
Large-scale structure studies using AGN in X-ray surveys - Challenges from XBOOTES and prospects for eROSITA

Alexander Kolodzig



München 2015

Large-scale structure studies using AGN in X-ray surveys - Challenges from XBOOTES and prospects for eROSITA

Alexander Kolodzig

Dissertation
an der Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Alexander Kolodzig
aus Berlin

München, den 17.03.2015

Erstgutachter: Prof. Rashid Sunyaev

Zweitgutachter: Prof. Werner Becker

Tag der mündlichen Prüfung: 13.04.2015

Die Vermessung des Universums mit aktiven galaktischen Kernen aus Röntgen-Himmelsdurchmusterungen Herausforderungen von XBOOTES und Perspektiven für eROSITA

von Alexander Kolodzig

Zusammenfassung

Großflächige Röntgen-Himmelsdurchmusterungen sind von großer Bedeutung für die Bestimmung großskaliger Strukturen im Universum. Aktuelle wissenschaftliche Erkenntnisse über diese Strukturen können mit Hilfe von aktiven galaktischen Kernen (engl. “active galactic nuclei”, AGN) einerseits durch die Durchführung von größeren und tieferen Röntgen-Beobachtungen und andererseits durch die Untersuchung der Intensitätsfluktuationen (engl. “surface brightness fluctuations”) der unaufgelösten kosmischen Röntgenhintergrundstrahlung (engl. “cosmic X-ray background”, CXB) verbessert werden.

In dem ersten Teil der vorliegenden Dissertation haben wir den möglichen wissenschaftlichen Erkenntnisgewinn durch die Untersuchung der großskaligen Strukturen unter der Verwendung von allen detektierbaren AGN in der zukünftigen, vollständigen Röntgen-Himmelsdurchmusterung von eROSITA (engl. “eROSITA all-sky survey”, eRASS) quantifiziert. Wir können zeigen, dass eRASS circa drei Millionen AGN im Energieband $0.5 - 2.0$ keV detektieren wird. Daraus ergibt sich, dass eRASS im Vergleich zur vorhergehenden Himmelsdurchmusterung von ROSAT (engl. “ROSAT all-sky survey”, RASS) einen ~ 30 mal größeren AGN Datensatz haben wird, weil es Objekte mit einem ~ 30 mal schwächeren Röntgenfluss detektieren wird. Desweiteren zeigen wir auf, dass dieser einzigartige AGN Datensatz eine mittlere Leuchtkraft (Median) von $\sim 10^{44}$ erg s $^{-1}$ haben wird, welche charakteristisch für die gesamte AGN-Population im betrachteten Energieband ist. Wir prognostizieren für den Datensatz eine mittlere Rotverschiebung (Median) von $z \approx 1$. Etwa 40 % dieser AGN befinden sich in dem Rotverschiebungsintervall von $z = 1 - 2$, in dem der größte Teil der Röntgenstrahlung von AGN erzeugt wird. Voraussichtlich $10^4 - 10^5$ der AGN werden eine Rotverschiebungen von $z \gtrsim 3$ haben und $2000 - 30\,000$ überschreiten die Rotverschiebung von $z = 4$, worunter sich auch voraussichtlich einige der kosmologisch jüngsten AGN befinden. Wir demonstrieren in dieser Arbeit, dass dieser

AGN Datensatz mit seinen einzigartigen Eigenschaften unseren Kenntnisstand über die räumliche Dichte von AGN als Funktion von Rotverschiebung und Leuchtkraft über eine weite kosmologische Zeitspanne stark verbessern kann. Desweiteren zeigen wir auf, dass dieser Datensatz uns eine detaillierte Studie der räumliche Korrelation zwischen der Dunklen Materie (engl. “Dark Matter”, DM) und AGN ermöglicht, weil wir zum ersten Mal relativ kleine Intervalle für die Rotverschiebung und Leuchtkraft bilden können. Die resultierenden Ergebnisse werden unseren Erkenntnisstand über das Wachstum supermassereicher schwarze Löcher über einen breiten kosmologischen Zeitraum und deren Auswirkungen auf die Galaxienevolution wesentlich verbessern. Wir zeigen zum erstem Mal auf, dass bedingt durch die Größe und Tiefe von eRASS, der AGN Datensatz mit Hilfe der Detektierung von baryonischen akustischen Oszillationen (engl. “baryon acoustic oscillations”, BAOs) als ein kosmologisches Lineal genutzt werden kann. Es wird möglich sein, BAOs mit hoher statistischer Signifikanz im bisher nicht vermessenen Rotverschiebungsbereich von $z \sim 1 - 2$ zu detektieren. Dies wird dazu beitragen die Parameter des kosmologischen Standardmodells genauer zu bestimmen.

In dem zweiten Teil dieser Dissertation haben wir die bisher genaueste Messung der Intensitätsfluktuationen des unaufgelösten CXB im Energieband $0.5 - 2.0$ keV für die Winkelskalen von $\lesssim 17'$ durchgeführt. Dafür nutzten wir die Röntgenbeobachtungen vom XBOOTES-Feld, welches die derzeit größte zusammenhängende Himmelsdurchmusterungen des Röntgenteleskops *Chandra* darstellt. Wir stellen fest, dass für kleine Winkelskalen ($\lesssim 2'$) das gemessene Leistungsspektrum der CXB Intensitätsfluktuationen weitestgehend im Einklang mit einem konventionellen AGN-Intensitätsfluktuations-Modell ist, wenngleich wir eine Abweichung von 30 % feststellen. Diese Abweichung kann allerdings auch als eine Herausforderung angesehen werden, um unseren bisherigen Kenntnisstand über die räumliche Korrelation zwischen DM und unaufgelösten AGN zu erweitern und komplexere AGN-Intensitätsfluktuations-Modelle mit unserer Messung zu testen. Für Winkelskalen von $\gtrsim 2'$ messen wir eine signifikante Abweichung von bis zu einer Größenordnung von dem beobachteten Leistungsspektrum im Vergleich zu dem AGN-Intensitätsfluktuations-Modell. Wir veranschaulichen, dass eine instrumentelle Ursache ausgeschlossen werden kann und zeigen auf, dass diese starke Abweichung im beobachteten Leistungsspektrum mit keiner bekannten Röntgenquelle beschrieben werden kann. Möglicherweise ist mehr als eine Röntgenquelle für die starke Abweichung verantwortlich. Allerdings sollte die dominierende Röntgenquelle, welche die starke Abweichung verursacht, von extragalaktischer Natur sein. Zum Abschluss geben wir Vorhersagen, in welcher Form eRASS die Studien des unaufgelösten CXB vorantreiben wird.

Large-scale structure studies using AGN in X-ray surveys

Challenges from XBOOTES and prospects for eROSITA

by Alexander Kolodzig

Abstract

Large X-ray surveys are a powerful tool to study the large-scale structure (LSS) of the Universe. The scientific impact of LSS studies using active galactic nuclei (AGN) in X-ray surveys can be significantly increased by conducting wider and deeper X-ray surveys and studying the surface brightness fluctuations of the unresolved cosmic X-ray background (CXB).

In the first part of this Thesis, we have investigated the prospects of using the AGN sample to be detected by the upcoming eROSITA all-sky survey (eRASS) for LSS studies. We show that eRASS will detect about 3 million AGN in the $0.5 - 2.0$ keV band. This will result in a ~ 30 times larger number of sources and a ~ 30 times better sensitivity than its ~ 25 year old predecessor, the ROSAT all-sky survey (RASS). We show that this unprecedented AGN sample will have a median luminosity of $\sim 10^{44}$ erg s $^{-1}$, which is typical for the entire AGN population in this energy band. It will have a median redshift of $z \approx 1$ and approximately 40 % of the objects will be in the redshift range of $z = 1 - 2$, where the bulk of the X-ray emission of AGN is produced. About $10^4 - 10^5$ AGN are predicted to be beyond redshift $z = 3$ and about 2000 – 30 000 beyond $z = 4$, which will potentially include some of the earliest AGN in the Universe. We demonstrate that, given these unique properties, the eRASS-AGN sample will be able to significantly improve our current knowledge of the AGN spatial density as a function of redshift and luminosity over a wide range of cosmic time. Further, we show that it will enable us, for the first time, to perform detailed redshift- and luminosity-resolved studies of the clustering strength of X-ray selected AGN. All these measurements will dramatically improve our understanding of the growth of supermassive black holes over cosmic time and its implications for galaxy evolution. We demonstrate for the first time that, given the breadth and depth of eRASS, it will be possible to use AGN as a cosmological probe via baryon acoustic oscillation (BAO) measurements. We will be able to convincingly detect BAOs in the currently uncharted redshift range of $z \sim 1 - 2$, which will improve the constraints on the cosmological model.

In the second part of this Thesis, we have conducted the most accurate measurement to date of the brightness fluctuations of the unresolved CXB in the $0.5-2.0$ keV band for angular scales of $\lesssim 17'$. For this we used the XBOOTES survey, the currently largest continuous survey of the X-ray telescope *Chandra*. We find that on small angular scales ($\lesssim 2'$) the observed power spectrum of the brightness fluctuations is broadly consistent with the conventional AGN clustering model, although with a 30% deviation. This deviation nevertheless presents a good opportunity to improve our understanding of clustering properties of unresolved AGN by testing more sophisticated clustering models with our measurement. For angular scales of $\gtrsim 2'$ we measure a significant excess with up to an order of magnitude difference in comparison to the standard AGN clustering model. We demonstrate that an instrumental origin can be excluded. However, we also show that the excess can neither be explained with any known X-ray source population by looking at strength of its clustering signal and the shape of its energy spectrum. It might be caused by more than one type of source but the dominant source appears to have extragalactic origin. Finally, we make predictions on how eRASS will be able to advance the studies of the unresolved CXB.

List of acronyms

2dFGRS Two Degree Field Galaxy Redshift Survey (optical survey)

2pCF 2-point correlation function

3D three-dimensional

A10 Aird et al. (2010) (reference for the XLF of AGN in the $2.0 - 10.0$ keV band)

ACIS-I Detector of *Chandra*: 2×2 array of four 1024×1024 pixel CCDs of ACIS (Advanced CCD Imaging Spectrometer)

AGN Active galactic nucleus/nuclei (Sect. 1.2)

APEC Astrophysical Plasma Emission Code: Spectral model in XSPEC for a collisionally-ionized diffuse gas (e.g. Sect. 4.3.2)

BAO Baryon acoustic oscillation (e.g. Sect. 1.6 & Sect. 3.4)

BH Black hole

BKG Background

CDF *Chandra* deep field (X-ray survey)

CDFN CDF north (X-ray survey)

CDFS CDF south (X-ray survey)

ChAMP *Chandra* Multi-wavelength Project (Serendipitous X-ray Survey using *Chandra* Archival Data)

Chandra X-ray satellite (Sect. 1.4)

- CIAO** *Chandra* Interactive Analysis of Observations (software to analyze *Chandra* data)
- CL** Confidence level
- CMB** Cosmic microwave background
- COSMOS** Cosmic Evolution Survey (2 deg² multi-wavelength survey)
- CXB** Cosmic X-ray background (Galactic and extragalactic emission, e.g Chapter 4)
- DE** Dark energy
- DM** Dark matter
- DMH** Dark matter halo
- eBOSS** Extended baryon oscillation spectroscopic survey (Optical galaxy survey to detected BAOs)
- eRASS** eROSITA all-sky survey (X-ray survey, e.g. Sect. 1.5)
- eROSITA** X-ray satellite (e.g. Sect. 1.4)
- FOV** Field-of-view
- G08** Georgakakis et al. (2008) (reference for the $\log N - \log S$ of X-ray point sources)
- H05** Hasinger et al. (2005) (reference for the XLF of AGN in the 0.5 – 2.0 keV band)
- H06** Hickox & Markevitch (2006)
- Hard band** 2.0 – 10.0 keV (energy band)
- HEW** Half-energy width, alias half-power diameter (HPD) [arcsec]
- HOD** Halo occupation distribution
- IGM** Intergalactic medium
- IR** Infrared (wavelengths/energy-band)
- ISM** Interstellar medium
- ICM** Intracluster medium

-
- K05** Kenter et al. (2005) (main reference for XBOOTES)
- K07** Kim et al. (2007) (reference for the $\log N - \log S$ of X-ray point sources)
- L12** Lehmer et al. (2012) (reference for the $\log N - \log S$ of AGN)
- LDDE** Luminosity-dependent density evolution (X-ray luminosity function model, Sect. 2.4)
- LSS** Large-scale structure
- MCMC** Markov chain Monte Carlo (simulation)
- Mpc** - Mega parsec (distance unit, $1 \text{ Mpc} = 10^6 \text{ pc} \approx 3.1 \times 10^{22} \text{ m}$)
- PSF** Point-spread-function (response of a telescope system to a point source)
- PSPC** Position Sensitive Proportional Counters (detector of ROSAT)
- QSO** Quasi-stellar object or quasar (type of AGN) (Sect. 1.2)
- RASS** ROSAT all-sky survey (X-ray survey, Truemper 1993; Voges et al. 1999)
- ROSAT** X-ray satellite (Sect. 1.4)
- RSD** Redshift space distortions
- S/N** Signal-to-noise ratio
- SB** eROSITA Science Book (Merloni et al. 2012)
- SDSS** Sloan Digital Sky Survey (optical survey)
- SMBH** supermassive black hole
- SRG** Spektrum-Roentgen-Gamma (satellite) (e.g. Sect. 1.5)
- Soft band** $0.5 - 2.0 \text{ keV}$ (energy band)
- XBOOTES** X-ray survey (Chapter 4)
- XLf** X-ray luminosity function (e.g. Sect. 2.4)
- XMM** XMM-Newton (X-ray satellite, Sect. 1.4)
- XSPEC** X-ray spectral-fitting program (Arnaud 1996)

Contents

Zusammenfassung	v
Abstract	vii
List of acronyms	ix
1 Introduction	1
1.1 Scientific Rationale	1
1.2 Active galactic nuclei	5
1.3 Large-scale structure studies with AGN	7
1.4 X-ray telescopes	11
1.5 The eROSITA all-sky survey	14
1.6 Baryon acoustic oscillations	15
1.7 Overview of the Thesis	18
 I Large-scale structure studies with AGN - prospects for eRosita	 19
2 Statistical properties of AGN in eRASS	21
2.1 Introduction	21
2.2 Sensitivity	23
2.2.1 Instrumental and cosmic background	23
2.2.2 Average exposure and sensitivity	26
2.2.3 Confusion limit	27
2.2.4 Sensitivity map	28
2.3 AGN number counts	31
2.4 X-ray luminosity function of AGN	33
2.5 Luminosity and redshift distribution of detected AGN	37

2.5.1	Luminosity distribution	41
2.5.2	Redshift distribution	41
2.5.3	Uncertainties	44
2.5.4	High-redshift AGN	48
2.6	Redshift determination with the iron $K\alpha$ line	50
2.7	Optical counterparts	53
2.8	Summary and conclusions	55
3	Large-scale structure studies with AGN of eRASS	57
3.1	Introduction	57
3.2	Angular power spectrum	58
3.2.1	Uncertainties	62
3.2.2	Results	62
3.3	Linear bias factor	63
3.3.1	Method	63
3.3.2	Results	67
3.4	Baryon acoustic oscillations	67
3.4.1	Method	68
3.4.2	Results	70
3.5	Discussion and conclusions	73
3.5.1	Linear bias factor	73
3.5.2	BAOs	74
3.5.2.1	Comparison with dedicated BAO surveys	75
3.5.3	Redshift data	79
3.6	Summary	80
II	Large-scale structure studies with the unresolved cosmic X-ray background - challenges from XBOOTES	83
4	Brightness fluctuations in the unresolved CXB of XBOOTES	85
4.1	Introduction	85
4.2	Data preparation and processing	87
4.2.1	Exposure map and mask	88
4.2.2	Removing resolved sources	89
4.2.2.1	Point sources	89
4.2.2.2	Extended sources	94
4.2.2.3	Summary	94
4.2.3	Removing background flares	94

4.2.4	Instrumental background and background-subtracted map . . .	95
4.2.5	Fluctuation maps	96
4.3	Components of the unresolved CXB	97
4.3.1	Instrumental-background model	97
4.3.2	Source model	100
4.3.3	Galactic emission	104
4.3.4	Extragalactic emission	104
4.3.4.1	Unresolved AGN population	107
4.4	Brightness fluctuations of the unresolved CXB	109
4.4.1	Formalism	110
4.4.2	PSF-smearing model	111
4.4.3	Photon shot noise	113
4.4.3.1	Estimators	114
4.4.3.2	Evaluation	117
4.4.4	Tests for systematics	121
4.4.4.1	Mask effect	121
4.4.4.2	Total-count map vs. background-subtracted map . . .	124
4.4.4.3	Instrumental, and stowed-background map	124
4.4.4.4	Vignetting	124
4.4.5	Power spectrum of the brightness fluctuations	127
4.4.6	Clustering signal of unresolved AGN	127
4.4.7	An excess at angular scales of $\gtrsim 2'$	132
4.4.7.1	Normal galaxies	132
4.4.7.2	Galaxy clusters & groups and WHIM	133
4.4.7.3	Energy spectrum of the excess	135
4.5	Discussion and summary	138
4.5.1	Unresolved AGN	139
4.5.2	The excess in the clustering signal	140
4.5.3	eROSITA forecast	141
	Conclusions	143
	References	147
	Acknowledgment	156

List of Figures

- 1.1 Distribution of major extragalactic X-ray surveys by Chandra (purple), XMM-Newton (blue), ROSAT (green), and Einstein (cyan) in the survey-average point-source sensitivity ($0.5 - 2.0$ keV, $[\text{erg cm}^{-2} \text{s}^{-1}]$) versus survey area plane. Encircled points mark continuous surveys. The vertical dotted line shows the solid angle for the entire sky. Also shown is the prediction for the upcoming eROSITA all-sky survey (eRASS) in red (see Sect. 1.5 and also Fig. 2.3). (from eROSITA Science Book, Merloni et al. 2012; also see Brandt & Hasinger 2005) 2
- 1.2 Spatial distribution of galaxies detected in the Two Degree Field Galaxy Redshift Survey (2dFGRS, <http://www.2dfgrs.net>, Colless et al. 2001). Galaxies are not randomly distributed. Instead they follow a hierarchical structure, which traces the underlying dark matter distribution. 8
- 1.3 Focusing mechanism of an X-ray telescopes on the example of XMM-Newton using grazing incidence optics with a combination of a hyperbolic and parabolic mirrors to make use of the very small reflection angle of X-rays. (from “XMM-Newton Users Handbook”, ESA, XMM-Newton SOC, Fig. 2). 12
- 1.4 First convincing detection of BAOs in the clustering signal of galaxies. *Top*: The power spectrum of 2dFGRS galaxies (Cole et al. 2005, Fig. 12). The BAOs are visible as an amplitude oscillation in the power spectrum. The lower panel shows the measured power spectrum (including BAOs) divided by a theoretical power spectrum without BAOs (black crosses). *Bottom*: The correlation function of SDSS galaxies (Eisenstein et al. 2005, Fig. 2) with different cosmological models. The BAOs are visible as the hump around $110 h^{-1}$ Mpc. 16

1.5	Comparison of recent BAO distance-redshift measurements from different galaxy surveys with cosmological models (Weinberg et al. 2013, Fig. 8). The quantity $D_V(z)$ combines the three-dimensional measurement of the acoustic scale from the line-of-sight (redshift space) and the transverse scale (Eisenstein et al. 2005, Eq. 2).	17
2.1	Average point-source detection sensitivity of the eROSITA telescope as a function of the exposure time in the soft (red lines, 0.5 – 2.0 keV) and hard (blue lines, 2.0 – 10.0 keV) energy bands for the extragalactic sky. The horizontal dashed lines are the confusion limits for one source per 40 telescope beams (PSF HEW). The vertical line on the left shows the exposure time for a single scan at 100 % observing efficiency, the other vertical lines indicate the average exposure times for different survey durations and at the ecliptic poles (for the soft-band confusion limit) at 80 % observing efficiency.	24
2.2	Four-year soft-band (0.5 – 2.0 keV) sensitivity map of eRASS in Galactic coordinates (l , b), based on the exposure time map of J. Robrade (priv. comm.) for a continuous Sun-pointing and based on the N_H -map of Kalberla et al. (2005). The two black horizontal curves enclose the Galactic plane ($ b < 10^\circ$), which is excluded from our calculation, and the two regions encircled by black curves are our defined ecliptic poles, where the exposure time was set to 20.0 ksec. The red vertical line in the horizontal color bar shows the average sensitivity (from Table 2.2).	29
2.3	The sensitivity – solid angle curves for the soft (<i>left panel</i>) and hard (<i>right</i>) bands. The vertical dashed lines show the corresponding average sensitivities from the Table 2.2. The horizontal line shows the solid angle of the extragalactic sky.	30
2.4	Cumulative number counts $N(> S)$ for observed X-ray point-sources in the soft (<i>left</i>) and hard band (<i>right</i>). The blue dashed-dotted curve is from Kim et al. (2007, Table 3, ChaMP+CDFs) and the red dashed curve from Georgakakis et al. (2008). The vertical solid lines indicate the point-source detection sensitivity for different survey durations assuming 80 % observing efficiency, and the sensitivity at the 20 ksec exposure time, corresponding to the confusion limit in the soft band (left-most line marked "Poles").	32

2.5	Number of AGN in different redshift and luminosity bins expected to be detected in the course of the four-year survey in the soft band ($0.5 - 2.0$ keV). White empty bins with no number correspond to zero sources. The dotted line corresponds to the detection limit of eRASS. In the soft-band plot, the numbers in brackets are for the XLF without the exponential redshift cutoff, they are given only if the difference exceeds 10 %.	38
2.6	Same as Fig. 2.5, but for the hard band ($2.0 - 10.0$ keV).	39
2.7	Differential (<i>left</i>) and cumulative (<i>right</i>) luminosity distribution in the soft (<i>top</i>) and the hard (<i>bottom</i>) band for the four-year eRASS AGN sample in the extragalactic sky (<i>black</i>) and at the ecliptic poles only (<i>red</i>). The distributions are normalized to unity to facilitate comparison of the shapes. The dotted curves in the top panels were computed without the high-redshift cutoff in the soft-band XLF (see Sect. 2.4). The dashed black vertical lines in the right panels show the luminosity corresponding to the 10 % fraction of sources.	40
2.8	Same as Fig. 2.7, but for the redshift distribution.	42
2.9	Differential luminosity (<i>top</i>) and redshift (<i>bottom</i>) distributions in the soft (<i>left</i>) and hard (<i>right</i>) bands for the entire extragalactic AGN sample after four years (<i>black</i>) and for the brightest 10 % (<i>blue</i>) and faintest 10 % (<i>red</i>). The black curves are same as in Figs. 2.7 and 2.8, but without renormalization to unity.	43
2.10	Number of eRASS AGN in the soft band ($0.5 - 2.0$ keV) as a function of the redshift for different luminosity groups in a sky area similar to that covered by SDSS ($14\,000$ deg ²). The solid and dashed histograms show predictions based on our default soft-band XLF with and without exponential high-redshift cutoff (Sect. 2.4), respectively.	45
2.11	Differential (<i>left</i>) and cumulative (<i>right</i>) luminosity (<i>top</i>) and redshift (<i>bottom</i>) distributions of the soft-band extragalactic sky sample computed using different XLFs. The thick solid black curves show predictions based on the default XLF model. Also shown are predictions for the XLF model of Miyaji et al. (2000, Table 3), Hasinger et al. (2005, Table 5, without the exponential redshift cutoff, dotted curve), and Ebrero et al. (2009). To facilitate the comparison of shapes, differential distributions in left panels are normalized to unity.	46

- 2.12 Numbers of high-redshift AGN, $N(> z)$, expected in the soft band ($0.5 - 2.0$ keV) for the extragalactic sky after four years. The thick solid black curve shows the prediction based on the default XLF model. Also shown are predictions for the XLF model of Miyaji et al. (2000, Table 3), Hasinger et al. (2005, Table 5, without the exponential redshift cutoff, dotted curve), and Ebrero et al. (2009). To obtain these curves we integrated the XLFs to the highest redshift of $z = 7$. All curves are rescaled to match the average source density computed with our default model (Sect. 2.3). 49
- 2.13 Feasibility of using the iron $K\alpha$ line for the redshift determination at $z = 0$. See Sect. 2.6 for details. *Top*: Fraction of catastrophic failures. *Middle*: Accuracy of the redshift determination. The black dashed line corresponds to the energy resolution of eROSITA. *Bottom*: Number of sources for which the redshift can be determined with the accuracy shown in the middle panel, the catastrophic failures excluded (black points show only the $\log N - \log S$). 51
- 2.14 Cumulative I -band AB-magnitude distribution of AGN in the COSMOS and XBOOTES fields with the $0.5 - 2.0$ keV flux exceeding the four-year eRASS detection threshold. The thickness of the curves represents the standard deviation of a binomial distribution. The vertical lines show the photometric sensitivities of the SDSS (21.3 mag) and Pan-STARRS PS1 (22.6 magrr) and the magnitude limit for SDSS spectroscopy (19.1 mag). 54
- 3.1 Angular power spectrum of the full eRASS AGN sample (after four years) in the soft band ($0.5 - 2.0$ keV) for the extragalactic sky ($f_{\text{sky}} \approx 0.83$) and $0 < z < 5$. The grayshaded area and the blue histogram show the 1σ uncertainty region (Eq. 3.4) without and with ℓ -binning, respectively. The horizontal dotted line shows the level of shot noise, which was already subtracted from the angular power spectrum. For multipoles above the vertical dashed line (representing $l_{\text{max}} \approx 500$) our assumption of a linear clustering starts to break down. Therefore, we did not consider these multipoles in our subsequent calculations. 59
- 3.2 Same as Fig. 3.1, but with the angular power spectra for various narrow and broad redshift ranges added (see Sect. 3.2.2). 60
- 3.3 Signal-to-noise ratio of the amplitude of the angular power spectrum (Eq. 3.5) as a function of the redshift for different sky fractions. A $\Delta z = 0.2$ binning is assumed. 64

3.4	Signal-to-noise ratio of the amplitude of the angular power spectrum as a function of the survey duration for the redshift bin $0.8 < z < 1.0$ at different sky fractions.	65
3.5	Same as Fig. 3.3, but for different luminosity ranges (in units of $\log(L[\text{erg s}^{-1}])$, $0.5 - 2.0 \text{ keV}$) and a sky coverage of $10\,000 \text{ deg}^2$	66
3.6	Baryon acoustic oscillations in the power spectrum for the extragalactic sky in the redshift range $0.0 < z < 3.0$. At wavenumbers above the vertical dashed line (corresponding to $0.2 h \text{ Mpc}^{-1}$) our assumption of a linear clustering starts to break down. The red curve shows the original input model for the BAOs.	69
3.7	Same as Fig. 3.6, but for different redshift ranges and sky coverages.	71
3.8	Confidence level of a BAO detection as a function of sky coverage for different redshift ranges (see Sect. 3.4.1 for more explanations). The vertical gray dashed line shows the area of the extragalactic sky. . . .	72
3.9	Effective volumes of BAO surveys listed in Table 3.1 as a function of the wavenumber. Effective volumes are computed for redshift ranges indicated in the plot.	77
3.10	Confidence level of the BAO detection in the full extragalactic sky as a function of the redshift slice thickness (see Sect. 3.4.1) for different redshift ranges.	81
4.1	Shape of the PSF for <i>Chandra</i> ACIS-I (averaged over all CCDs) for different offset angles (θ) for the source flux of $0.63 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$ ($0.5 - 2.0 \text{ keV}$). The vertical dashed line shows the radius of the circular exclusion area for this flux group (Table 4.1). The horizontal dotted lines show levels corresponding to 100 % and 10 % of the surface brightness of unresolved AGN (Table 4.7).	90
4.2	Average total counts of the unresolved CXB ($0.5 - 2.0 \text{ keV}$) per XBOOTES observation as a function of radius of the circular exclusion area of point sources. Black crosses: using different radius intervals for the first three flux groups of Table 4.1 (see Text). The average radius was computed based on the number of point sources per flux group. Gray crosses: using the same radius interval for the first three flux groups. The radius of the fourth flux groups is fixed to the value in Table 4.1. Green dotted line: Average radius for the definition in Table 4.1. Red dashed line: radius, where the total removed area is equal to half the area of the ACIS-I FOV.	91

4.3	Energy spectrum of the ACIS-I stowed background to define our instrumental background model. Black crosses: data points with one standard deviation as error-bars. Green Curve: Total background model. Dotted Curves: single components of the model (see Table 4.2).	98
4.4	The stacked energy spectrum of the unresolved CXB from 118 XBOOTES observations (black crosses) for the 0.5 – 2.0 keV band, which we use for our brightness fluctuation analysis (Sect. 4.4), together with the ratio of model and data. Green solid curve: Total model (Source and background model). Dark blue solid curve: Source model (Table 4.4). Light blue dotted curve: absorbed powerlaw model (<code>phabs(powerlaw)</code>). Orange dotted curve: <code>APEC</code> (collisionally-ionized diffuse gas) model. Pink solid curve: Instrumental background model (Sect. 4.3.1). Black dotted curves: components of the instrumental background model. .	102
4.5	Same as Fig. 4.4 for the entire energy rang, which we use for the spectral fit.	103
4.6	Differential flux distribution (alias the flux production rate per solid angle, Eq. 4.6-4.7) for the unresolved AGN in the XBOOTES survey in the 0.5 – 2.0 keV band as a function of redshift (<i>top</i>) and luminosity (<i>bottom</i>). The dashed lines show the median.	108
4.7	<i>Two upper panels:</i> Average power spectrum of the PSF of <i>Chandra</i> ACIS-I (gray crosses, average of all offset angles and CCDs) and our empirical PSF-smearing model (red curve, Eq. 4.15). In the first panel we also show the average PSF power spectrum (dotted curves) for different offset angles from 0' (black) to 11' (red), where the thickness of the curve indicates the weight for the weighted average over all offset angles. <i>Lower panel:</i> Ratio between the measurement and our PSF-smearing model.	112
4.8	Shot-noise-subtracted power spectrum $P(k)$ (Eq. 4.12) in the 9.5 – 12.0 keV band. Red crosses (all panels): using the analytical shot-noise estimate (default) P_{Shot} (Eq. 4.16). Blue crosses (upper panels): using the high-frequency based shot-noise estimate $P_{\text{Shot}}^{(\text{HF})}$ (Sect. 4.4.3.1). Green crosses (lower panels): using the observation-splitting based estimate $P_{\text{Shot}}^{(\text{OS})}$ (Sect. 4.4.3.1). Colored error-bars: Standard deviation due to error propagation of all binned frequencies. Black error-bars: Standard deviation of the sample mean of all binned frequencies. Vertical blue dashed line: lower limit $k_{\text{min}}^{(\text{HF})}$ to estimate $P_{\text{Shot}}^{(\text{HF})}$	118

- 4.9 Same as Fig. 4.8 for the $0.5 - 2.0$ keV band. Purple solid curve: AGN clustering model (Sect. 4.4.6), which is multiplied with our PSF-smearing model (Sect. 4.4.2). We see that our PSF-smearing model can describe the measured power spectrum up to high frequencies. 120
- 4.10 To demonstrate the mask effect we use an ad-hoc model (purple curve), which represents our measured power spectrum ($0.5-2.0$ keV, Fig. 4.14), and convolve it with the full mask (red crosses, regions of resolved sources masked out) and only with the ACIS-I FOV-Mask (green crosses, regions of resolved sources *not* masked out) of all XBOOTES observations. Also shown our AGN clustering model (blue curve, Sect. 4.4.6). 122
- 4.11 Shot-noise-subtracted power spectrum $P(k)$ (Eq. 4.12) in the $0.5 - 2.0$ keV band computed with the total-count map $\mathbf{C}_X^{\text{Obs.}}$ (red crosses) and the background-subtracted map $\mathbf{C}_X^{\text{Scr.}}$ (green crosses, Eq. 4.3, Sect. 4.2.4). Purple solid curve: AGN clustering model (Sect. 4.4.6) multiplied with our PSF-smearing model (Sect. 4.4.2). We can see that for the shown frequencies range, there is no significant difference in computing the power spectrum with those two maps. 123
- 4.12 Flux-normalized, shot-noise-subtracted power spectra $Q(k)$ (Eq. 4.14) of the total-count map $\mathbf{C}^{\text{Obs.}}$ (red crosses) and of the ACIS-I stowed-background map $\mathbf{C}^{\text{sto.BKG}}$ (blue crosses, Sect. 4.2.4) for the indicated energy bands, and for the total-count map $\mathbf{C}^{\text{Obs.}}$ for the $9.5 - 12.0$ keV band (green crosses). 125
- 4.13 Shot-noise-subtracted power spectra $P(k)$ (Eq. 4.12) of the total-count map in the case of using the original (red crosses) and average (blue crosses) exposure map (Sect. 4.2.1). 126
- 4.14 The power spectrum of the brightness fluctuations of the unresolved CXB of XBOOTES in the $0.5 - 2.0$ keV band. Blue curve: AGN clustering model (Sect. 4.4.6, including the PSF-smearing model) with a $\pm 30\%$ envelope. 128
- 4.15 The power spectrum of the brightness fluctuations of the unresolved CXB of XBOOTES in the $1.0 - 2.0$ keV band. Blue curve: AGN clustering model (Sect. 4.4.6, including the PSF-smearing model) with a $\pm 30\%$ envelope. 129

-
- 4.16 Average power spectrum of the excess (black crosses) as a function of energy computed with Eq. (4.21) together with several normalized spectral models. The gray crosses in the upper panel show the average power spectrum for the frequency interval $[0.2, 0.6[\times 10^{-2} \text{ arcsec}^{-1}$ (first term in Eq. 4.21). 136
- 4.17 Predicted differential flux distribution (alias the flux production rate per solid angle, Eq. 4.6-4.7) for the unresolved AGN in the eROSITA all-sky survey in the $0.5 - 2.0 \text{ keV}$ band as a function of redshift (*top*) and luminosity (*bottom*), assuming an average point-source sensitivity of $S_{0.5-2.0 \text{ keV}} = 1.1 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$. The dashed lines show the median. 142

List of Tables

2.1	Predicted background count rates.	22
2.2	Characteristic average parameters of the eROSITA all-sky survey. . .	23
2.3	Parameters of the LDDE model used to compute the luminosity and redshift distributions of the detected AGN.	35
2.4	Peak and median values of the redshift distribution of eRASS AGN .	42
3.1	Parameters of BAO surveys.	76
4.1	Radius of the circular exclusion area of resolved point sources in XBOOTES for different flux groups.	92
4.2	Parameter values of the instrumental background model in the spec- tral fit of the unresolved CXB of XBOOTES (Sect. 4.3, Fig. 4.4-4.5) .	99
4.3	Surface brightness of the instrumental background of XBOOTES in the 0.5 – 2.0 keV band [cts s ⁻¹ deg ⁻²].	99
4.4	Best-fit parameters of our spectral model (APEC + phabs(powerlaw)) of the unresolved CXB of XBOOTES (Fig. 4.4-4.5).	101
4.5	Surface brightness of the unresolved CXB of XBOOTES in the 0.5 – 2.0 keV band [10 ⁻¹² erg cm ⁻² s ⁻¹ deg ⁻²].	101
4.6	Extragalactic emission of XBOOTES [10 ⁻¹² erg cm ⁻² s ⁻¹ deg ⁻²]. . . .	105
4.7	Individual components of the unresolved extragalactic emission of XBOOTES in the 0.5 – 2.0 keV band [10 ⁻¹² erg cm ⁻² s ⁻¹ deg ⁻²]. . . .	106
4.8	Parameters of our PSF model (Eq. 4.15)	113

Chapter 1

Introduction

1.1 Scientific Rationale

Since the beginning of this millennium there has been an increasing number of deep X-ray surveys, thanks to the high angular resolution and sensitivity of the X-ray telescopes *Chandra*¹ and *XMM-Newton*² (see Sect. 1.4 for details about X-ray telescopes). These new surveys were driven by the ambition to resolve and better understand the extragalactic components of the cosmic X-ray background (CXB³) (e.g. Brandt & Hasinger 2005; Alexander et al. 2013). This led to an ongoing increase in survey depth in order to resolve fainter objects and to detect objects at earlier cosmic times, which was accompanied with a decrease in survey area, as we can see in Fig. 1.1. As a comparison, the last all-sky survey (RASS) in the 0.5–2.0 keV band was performed by ROSAT⁴ (Truemper 1993; Voges et al. 1999) 25 years ago, which had an average point-source sensitivity of $\sim 2 \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$. The 4 Megasecond Chandra Deep Field-South (CDF-S), one of the deepest X-ray surveys to date, achieves a point-source sensitivity of $S_{0.5-2.0 \text{ keV}} \approx 9 \times 10^{-18} \text{ erg cm}^{-2} \text{ s}^{-1}$, but only has a sky coverage of $\sim 0.1 \text{ deg}^2$ (Xue et al. 2011). Such pencil-beam-like X-ray surveys are able to resolve more than 90 % of the CXB. Thanks to these deep surveys, we know now for certain that most of the CXB emission is caused by active galactic nuclei (AGN) (e.g. Comastri et al. 1995; Moretti et al. 2003; Hickox & Markevitch 2006,

¹<http://chandra.harvard.edu>

²<http://sci.esa.int/xmm-newton>

³ In the framework of this Thesis, the CXB consists of Galactic and extragalactic emission. We note that in the literature the term “CXB” is also sometimes used exclusively for extragalactic emission.

⁴<http://www2011.mpe.mpg.de/xray/wave/rosat>

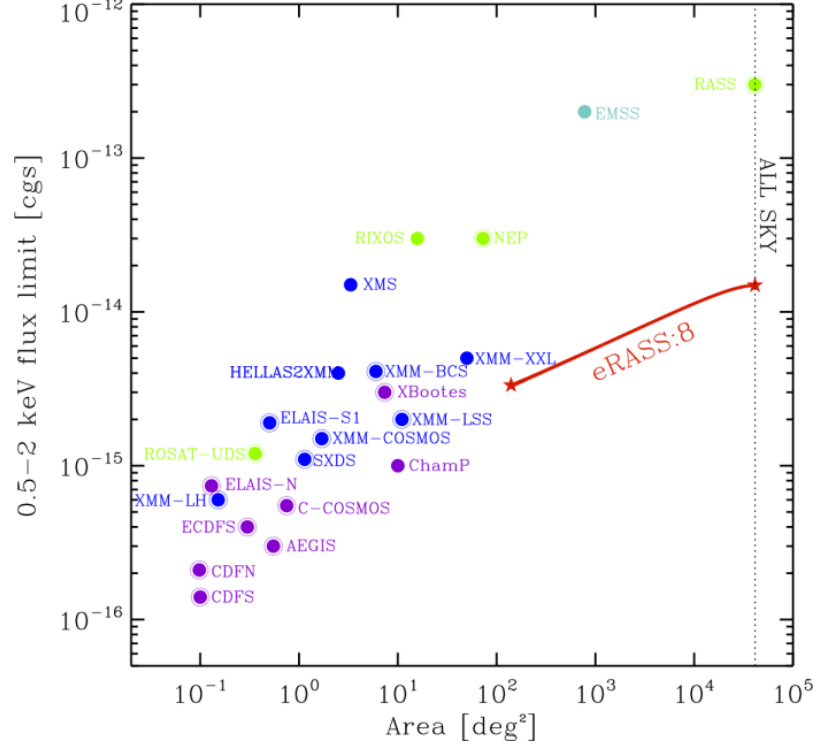


Figure 1.1: Distribution of major extragalactic X-ray surveys by Chandra (purple), XMM-Newton (blue), ROSAT (green), and Einstein (cyan) in the survey-average point-source sensitivity ($0.5 - 2.0$ keV, $[\text{erg cm}^{-2} \text{s}^{-1}]$) versus survey area plane. Encircled points mark continuous surveys. The vertical dotted line shows the solid angle for the entire sky. Also shown is the prediction for the upcoming eROSITA all-sky survey (eRASS) in red (see Sect. 1.5 and also Fig. 2.3). (from eROSITA Science Book, Merloni et al. 2012; also see Brandt & Hasinger 2005)

2007; Gilli et al. 2007; Moretti et al. 2012), which are accreting supermassive black holes (SMBHs) in the centers of some galaxies (see Sect. 1.2 for details about AGN). The combination of these deep X-ray surveys with more shallow ones and surveys in other wavelength bands made it possible to measure the spatial density of AGN to an unprecedented detail (e.g. Ueda et al. 2003; Hasinger et al. 2005; Aird et al. 2010; Ueda et al. 2014; Miyaji et al. 2015). This led to a significant improvement in our understanding of the evolution of AGN and the complex picture of galaxy evolution (e.g. Brandt & Hasinger 2005; Hopkins et al. 2006; Hickox et al. 2009; Alexander & Hickox 2012).

Thanks to the X-ray surveys of *Chandra*, *XMM-Newton*, and also the all-sky survey RASS, complemented by spectroscopic redshift surveys of sufficient size and depth, the field of clustering measurements with AGN developed into an important branch of large-scale structure (LSS) studies during the last decade (e.g. Cappelluti et al. 2012a; Krumpe et al. 2014). These measurements enable us to study in a statistical manner the triggering mechanisms of AGN activity and allow us to understand how these mechanisms depend on the AGN environment, such as the properties of their host galaxies and dark matter halos (DMHs), in which the galaxies reside (see Sect. 1.3 for more details about LSS studies with AGN). With this we can better understand how SMBHs co-evolve with their host galaxy and their DMH over cosmic time, which are essential questions in the field of galaxy evolution (e.g. Brandt & Hasinger 2005; Hopkins et al. 2006; Hickox et al. 2009; Alexander & Hickox 2012).

Despite the current success of LSS studies with AGN, there are still many open questions (see Sect. 1.3). These questions can not be answered to a statistically significant degree with the AGN detected in the current ensemble of X-ray surveys. These surveys are either too shallow in sensitivity (e.g. RASS) or too small in survey area (e.g. CDF-S), as we can see in Fig. 1.1. In this Figure, we can also see that the upcoming eROSITA all-sky survey (eRASS, red curve, Sect. 1.5) will fill this gap in the survey sensitivity versus survey area plane. eRASS will outperform in terms of survey area any continuous X-ray survey of *XMM-Newton* and *Chandra* by several orders of magnitude (Merloni et al. 2012). What consequences eRASS will have for the field of LSS studies with AGN is the focus of the first part of this Thesis (Chapter 2 and 3).

Another feature of almost all current studies of the clustering properties of AGN, and its implications for AGN and galaxy evolution, is that they are based on LSS studies with *resolved* AGN. This makes our knowledge biased towards bright and luminous AGN ($L_{0.5-2.0\text{ keV}} \gtrsim 10^{42} \text{ erg s}^{-1}$), in particular for higher redshifts ($z > 0.5$), due to the signal-to-noise ratio (S/N) cut for the spectroscopic redshift and the luminosity cut from the AGN identification process (Sect. 1.2) (e.g. Allevato et al.

2011, 2012, 2014; Krumpe et al. 2010b, 2012; Miyaji et al. 2011; Mountrichas & Georgakakis 2012; Mountrichas et al. 2013; Koutoulidis et al. 2013). However, an important question to ask is if we are able to extrapolate these clustering properties to less luminous AGN. These AGN trace galaxies at an earlier evolutionary stage with a less massive SMBH and/or smaller accretion rate than luminous AGN. A significant step towards answering this question is to study the surface brightness fluctuations of the unresolved CXB. This type of clustering measurement with the high angular resolution of *Chandra* has only been performed so far with very small surveys ($\sim 0.1 \text{ deg}^2$) (e.g. Cappelluti et al. 2012b, 2013; Helgason et al. 2014). In the second part of this Thesis (Chapter 4), we conduct the most accurate measurement to date of the brightness fluctuations of the unresolved CXB for angular scales of $\sim 3'' - 17'$. This is accomplished by using the XBOOTES survey, currently the largest continuous survey using *Chandra* (see Fig. 1.1). We show how this measurement is able to challenge our current understanding of the clustering properties of AGN and other X-ray source populations.

We can conclude that in order to further increase the scientific impact of LSS studies with AGN in the future, two major steps are necessary:

- Wider and deeper X-ray surveys have to be performed to build larger AGN samples and to increase the overall S/N.
- LSS studies with the unresolved CXB have to be further developed and utilized to fully exploit current and future X-ray survey data and to minimize the bias towards bright and luminous AGN.

Both steps lead to new measurements, which will challenge our current understanding of AGN clustering and its implications, and give us access to new scientific topics. The first step is addressed in the first part of this Thesis (Chapter 2 and 3), while we deal with the second step in the second part of this Thesis (Chapter 4).

In remaining part of this Chapter, we describe the major characteristics of AGN in the X-ray regime in Sect. 1.2. This is followed by an brief overview in Sect. 1.3 of the current results and problems of LSS studies with AGN. In Sect. 1.4 we describe the general concept of X-ray telescopes before we give a general introduction in Sect. 1.5 of eRASS. Since baryon acoustic oscillations (BAOs) are an important aspect of our work presented in Chapter 3, we give a more general description of them in Sect. 1.6. In the last section (1.7) of this Chapter, we give an overview of the structure of this Thesis.

1.2 Active galactic nuclei

It is now well established that most galaxies host a SMBH ($\gtrsim 10^6 M_\odot$) in their centers (e.g. Ferrarese & Ford 2005; Heckman & Best 2014). These SMBHs grow via one or more stages of intense matter accretion. During such a stage the nucleus of an galaxy becomes one of the most luminous objects in the Universe, known as the “Active Galactic Nucleus” or shorthand AGN⁵.

AGN have a high radiation output over a wide spectrum of energies spanning from the radio to the γ -ray regime. Due to the complexity of the energy spectrum, which changes with the accretion rate (or more precisely with the *Eddington ratio*) of the SMBH, with the properties of the host galaxy, and with the orientation towards the observer, and its strong variability over different timescales, there exists a *zoo* of different classifications for AGN (for an overview see e.g. Kembhavi & Narlikar 2000). This is a result of the discovery of spectral features of AGN in different wavelength bands. These features were eventually explained in an unified scheme, which is comprehensively described in Antonucci (1993) and Urry & Padovani (1995) (also see Heckman & Best 2014 for a more recent review). In framework of this Thesis, we will focus on the emission in the X-ray regime (0.5 – 10 keV).

The accreting SMBH of an AGN is surrounded by an extended accretion disk. The inner region of this disk can be described as a geometrically thin, optically thick, viscous accretion disk (Shakura & Sunyaev 1973; Pringle 1981), surrounded by a hot corona (e.g. Haardt et al. 1994). This is the region where most of the radiation power from the accretion process is released, and where most of the X-ray emission originates. Therefore, X-rays gives us the most direct view onto an accreting SMBH in comparison to radiation in lower wavelength bands, which originate at larger distances from the SMBH.

We can describe the X-ray spectrum of an AGN to a first order approximation with a *powerlaw*: $F(E) \propto E^{-\Gamma}$, where E is the photon energy and Γ the photon index. The photon index typically has a value of $\Gamma \sim 1.9$ with a typical dispersion of $\Delta\Gamma \sim 0.4$ (e.g. Reeves & Turner 2000; Gilli et al. 2007; Tozzi et al. 2006). The actual X-ray spectrum is much more complex and also variable over time (e.g. Mushotzky et al. 1993), but a powerlaw is a sufficient description for statistical studies with a large sample of AGN, such as studies of the number density (e.g. Lehmer et al. 2012), the spatial density (e.g. Ueda et al. 2014), or LSS studies (e.g. Allevato et al. 2011). The X-ray emission is absorbed along the line-of-sight by the gas of our Galaxy, which reduces the X-ray flux, particularly at low energies ($\lesssim 1$ keV). The absorption

⁵We use the same abbreviation also for the plural “Active Galactic Nuclei” throughout this Thesis.

strength depends on the hydrogen column density N_{H} in the line-of-sight, which has a typical value for the extragalactic sky ($|b| > 10^\circ$) of $N_{\text{H}} \approx 6 \times 10^{20} \text{ cm}^{-2}$ (arithmetic mean of the N_{H} -map of Kalberla et al. 2005). An example for an absorbed powerlaw can be seen in Fig. 4.4⁶ (light blue dotted curve). However, depending on orientation of the accretion disk of the SMBH towards the observer, the X-ray emission of an AGN can already be heavily absorbed within the host galaxy due to a torus of molecular gas and dust surrounding the accretion disk. This causes the AGN to be partially or completely obscured in the 0.5–2.0 keV band. Therefore, X-ray selected AGN are commonly separated into type-1 and type-2, where the former describes *unobscured* AGN ($N_{\text{H}} \lesssim 10^{21} \text{ cm}^{-2}$) and the latter *obscured* AGN ($N_{\text{H}} \gtrsim 10^{21} \text{ cm}^{-2}$) but the exact definition can vary between different studies (e.g. Gilli et al. 2007).

The great advantage of detecting AGN with X-rays is that it is the least ambiguous method to identify them (e.g. Alexander & Hickox 2012). First of all, AGN are the most numerous source of the CXB down to fluxes of $F_{0.5-2.0 \text{ keV}} \sim 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. There are only a few other types of extragalactic X-ray sources (Sect. 1.4), such as galaxy clusters & groups and accreting stellar binary systems (X-ray binaries) in normal galaxies. Secondly, AGN and galaxy clusters are the only X-ray sources with a persistent luminosity of $L_{0.5-2.0 \text{ keV}} \gtrsim 10^{42} \text{ erg s}^{-1}$. Given a sufficient angular resolution of the X-ray telescope (Sect. 1.4), most galaxy clusters can be detected as extended sources while AGN are detected as point sources. For lower luminosities, AGN identification becomes more difficult because other sources, such as galaxy groups and star-forming galaxies, have to be taken into account. Therefore, most X-ray selected AGN samples have a luminosity-cut at $L_{0.5-2.0 \text{ keV}} \approx 10^{42} \text{ erg s}^{-1}$ in order to avoid contamination by other types of sources. This is particularly necessary for higher redshifts ($z \gtrsim 0.5$), where source identification becomes generally more difficult and redshift information is generally less certain. This results in the effect that all derived quantities, such as the XLF or the clustering properties of AGN, from such AGN samples are biased towards X-ray bright and luminous AGN (e.g. Hasinger et al. 2005; Allevato et al. 2011; Krumpe et al. 2012). Since the bulk of X-ray selected AGN has a luminosity around $L_{0.5-2.0 \text{ keV}} \sim 10^{44} \text{ erg s}^{-1}$, this is not an issue for an average description of AGN. However, low-luminous AGN are also an interesting class of objects, because they trace galaxies at an earlier evolutionary stage with a less massive SMBH and/or smaller accretion rate than luminous AGN (e.g. Hickox et al. 2009). It remains to be seen, if the derived properties from AGN of $L_{0.5-2.0 \text{ keV}} \gtrsim 10^{42} \text{ erg s}^{-1}$ can be extrapolated to less luminous AGN.

⁶ Note that this is for demonstration purposes only, because the spectrum shown actually describes the entire extragalactic emission of the unresolved CXB, which contains more than just AGN emission. See also Table 4.7.

AGN gives us an unique way to study the growth of SMBHs over a wide range of cosmic time. There is clear evidence that the evolution of a galaxy is correlated with the growth of its SMBH, the strongest one being the tight correlation of the SMBH mass with the stellar velocity dispersion in the bulge of the galaxy (e.g. Ferrarese & Merritt 2000; Heckman & Best 2014). The exact mechanisms for such a correlation, such as *AGN feedback*, are still debated (e.g. Alexander & Hickox 2012). Nevertheless, since we know that almost every galaxy undergoes a phase of AGN activity, this makes AGN an essential tool for investigating galaxy evolution (e.g. Hopkins et al. 2006; Hickox et al. 2009). Thanks to the many deep and shallow surveys of the last decades in combination with surveys of other wavelength, it was possible to measure the spatial density of AGN as a function of redshift and luminosity to an unprecedented detail (e.g. Ueda et al. 2003; Hasinger et al. 2005; Aird et al. 2010; Ueda et al. 2014; Miyaji et al. 2015). These measurements established the luminosity-dependent density evolution (LDDE) model as the best description of the X-ray luminosity function (XLF) of AGN. They also led to a significant improvement in our understanding of the evolution of AGN and the growth of their SMBHs up to high redshift ($z \sim 3 - 5$). Studies with AGN XLF have shown that the peak in the AGN spatial density moves from earlier to the later cosmic time with decreasing X-ray luminosity. This is interpreted as an anti-hierarchical growth of SMBHs, where the high-mass end of SMBHs ($\sim 10^{7.5} - 10^9 M_\odot$) grow very efficiently at high redshift ($z \sim 1 - 3$) while the low-mass end ($\sim 10^6 - 10^{7.5} M_\odot$) have its largest growth rate at smaller redshift ($z \lesssim 1.5$) (e.g. Brandt & Hasinger 2005). These are essential results, which need to be considered for LSS studies with AGN (Sect. 1.3) and for understanding the evolution of galaxies in general (e.g. Hopkins et al. 2006; Hickox et al. 2009; Alexander & Hickox 2012).

1.3 Large-scale structure studies with AGN

LSS studies with X-ray surveys are a powerful tool to improve our understanding of the growth of SMBHs over cosmic time and its implications for galaxy evolution. Further, they can improve our constraints of the standard model of cosmology. The basic idea behind LSS studies is that galaxies trace the underlying dark matter (DM) distribution. By measuring the spatial distribution of galaxies we are able to map the three-dimensional (3D) structure of the Universe. One of the pioneering works in this field was performed with the Two Degree Field Galaxy Redshift Survey⁷ (2dFGRS,

⁷<http://www.2dfgrs.net>

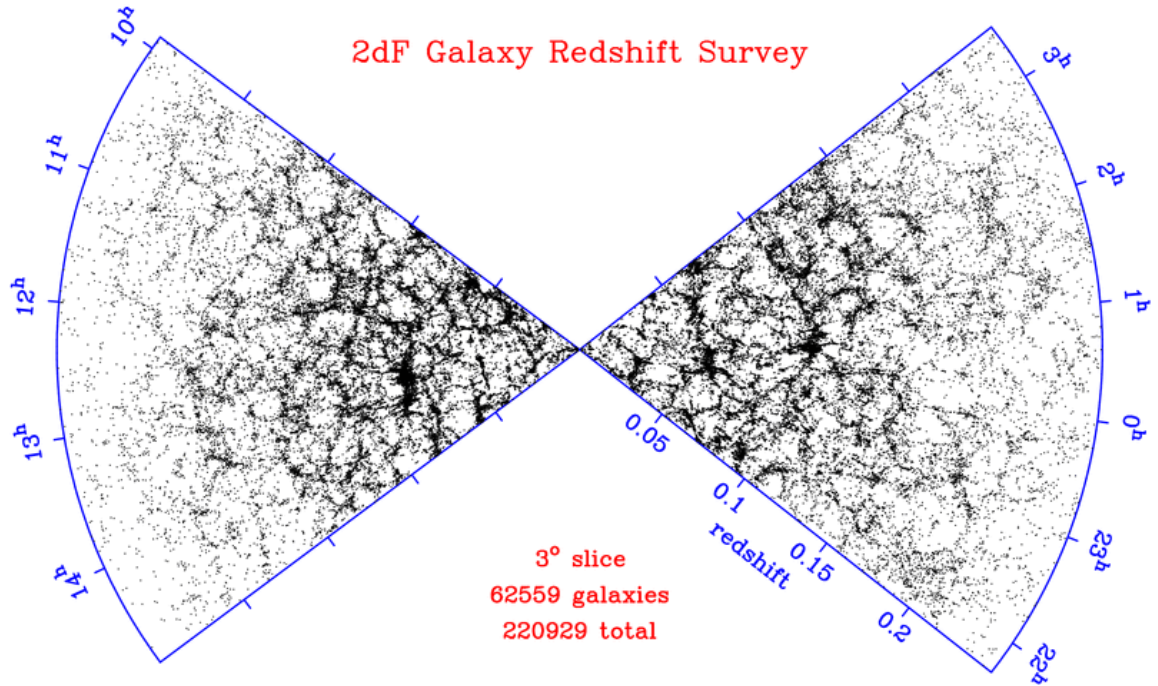


Figure 1.2: Spatial distribution of galaxies detected in the Two Degree Field Galaxy Redshift Survey (2dFGRS, <http://www.2dfgrs.net>, Colless et al. 2001). Galaxies are not randomly distributed. Instead they follow a hierarchical structure, which traces the underlying dark matter distribution.

Colless et al. 2001) and the Sloan Digital Sky Survey⁸ (SDSS, e.g. Alam et al. 2015). These optical surveys gave us two of the most precise maps of the 3D structure of the nearby Universe ($z \lesssim 0.5$), as we can see for 2dFGRS in Fig. 1.2. Their clustering measurements had a major impact in understanding of the LSS formation and put strong constraints on the standard cosmological model by convincingly detecting for the first time BAOs (see Sect. 1.6, Cole et al. 2005; Eisenstein et al. 2005; Hütsi 2006; Tegmark et al. 2006).

Since AGN are one of the most luminous type of objects in the Universe, we can detect them at much higher redshift than normal galaxies⁹. This makes AGN an ideal tracer for the LSS studies at a much earlier time of the Universe ($z \sim 1 - 3$), where normal galaxies are too faint to be used for LSS studies. However, since AGN are two to three orders of magnitude more sparse than normal galaxies, LSS studies with AGN require a much larger survey area than LSS studies with normal galaxies in order to achieve the same statistical significance in the clustering measurement (e.g. Cappelluti et al. 2012a). Thanks to the deep and wide X-ray surveys, such as COSMOS¹⁰ (Elvis et al. 2009; Hasinger et al. 2007) and RASS (Truemper 1993; Voges et al. 1999), complemented by the necessary spectroscopic redshift data, clustering measurements with X-ray selected AGN developed into an important branch of LSS studies during the last decade (e.g. Cappelluti et al. 2012a; Krumpe et al. 2014). With AGN clustering measurements we are able to constrain different triggering and fueling mechanisms for AGN as a function of AGN environment, such as properties of the host galaxies and DMH. This enables us to better understand how SMBHs co-evolve with their host galaxy and their DMH over cosmic time (e.g. Hickox et al. 2009; Fanidakis et al. 2013).

One of the important results from these studies is that X-ray selected AGN trace DMHs with the mass of $\sim 10^{13} h^{-1} M_{\odot}$ up to the current observational limit of $z \sim 3$ (e.g. Allevato et al. 2011; Krumpe et al. 2012; Mountrichas et al. 2013). This indicates that galaxy group-sized DMHs are the optimal environment to trigger AGN activity. However, it also appears that there is some tension with the results of clustering studies with optical selected AGN (quasars), which trace DMHs with a significantly smaller mass of $\sim 10^{12} h^{-1} M_{\odot}$. One interpretation is that X-ray selected and optical selected AGN are triggered by different mechanisms. In principle, this would be consistent with current theories of galaxy evolution (e.g. Hickox et al. 2009) and recent simulations (e.g. Draper & Ballantyne 2012; Hirschmann et al. 2012; Fanidakis

⁸<http://www.sdss.org>

⁹Hereby, we use the term “normal galaxies” to denote galaxies, which do not exhibit any detectable sign of AGN activity.

¹⁰<http://cosmos.astro.caltech.edu>

et al. 2012, 2013). However, there are also valid concerns about the significance of the discrepancy of the DMH mass between the X-ray selected and optical selected AGN. One alternative explanation can be that this difference is caused by a selection effect when creating the AGN samples (luminosity-cut in X-rays and color/spectral cut in optical) (e.g. Mendez et al. 2013). Another alternative explanation is that this difference arises due the use of different methods to estimate the DMH mass in the different studies (e.g Krumpe et al. 2014). Thereby, both alternatives are not mutually exclusive. In order to solve this issue, one needs to perform statistically meaningful comparisons, where similar selection criteria and the same luminosity-bins are used. Due to the small samples of X-ray selected AGN, this is currently not possible (e.g. Allevato et al. 2011; Krumpe et al. 2012; Mountrichas et al. 2013; Krumpe et al. 2014).

Another interesting result is that there is an indication of a weak luminosity dependence of the DMH mass from clustering studies of X-ray selected AGN (e.g. Krumpe et al. 2012; Mountrichas et al. 2013; Fanidakis et al. 2013). Such a dependence is in fact suggested by certain triggering/fueling mechanism for AGN, based on semi-analytic simulations (e.g. Fanidakis et al. 2012). The dominant fueling mechanism for AGN is cold gas accretion, which is either triggered by a merger of galaxies (*major/minor mergers*) or by a disk instability in the galaxy itself. Since this fueling mechanism is accompanied by a high star-formation rate within the host galaxy, is also called the *starburst mode*. This mode shows no strong dependence between the AGN luminosity and the DMH mass in semi-analytic simulations. In comparison, the *hot-halo mode* shows in fact a strong dependence between the AGN luminosity and the DMH mass in the same simulations (e.g. Fanidakis et al. 2012; Mountrichas et al. 2013; Fanidakis et al. 2013). Hereby, the SMBH is fueled by the hydrostatic hot gas of the DMH. In this accretion process, the AGN luminosity is tuned to the cooling rate of the gas, which is defined by the mass of the DMH. Unfortunately, a comparison between semi-analytic simulations and current observations did not lead to any definitive conclusions due to the insufficient S/N of observational data. In order to separate clearly these different AGN triggering modes, one needs to have a more detailed redshift- and luminosity-resolved clustering measurement of X-ray selected AGN and statistically significant clustering measurements for luminosities of $L_{0.5-2.0\text{ keV}} > 10^{44} \text{ erg s}^{-1}$.

The results, that X-ray selected AGN reside in DMHs with mass of $\sim 10^{13} h^{-1} M_{\odot}$ and that there are indications for a weak luminosity dependence of the DMH mass, are based on clustering measurement on large spatial scales ($r \gtrsim 2 h^{-1} \text{ Mpc}$), where one measures the clustering strength of AGN pairs from two different DMHs (*two-halo-term*). It is also interesting to measure the clustering strength of AGN pairs

within the same DMH (*one-halo-term*). This gives additional constraints on triggering/fueling mechanisms. However, measuring the one-halo-term is difficult, because of the small spatial scales ($r \lesssim 2 h^{-1}$ Mpc) and the general weak clustering signal of the one-halo-term of AGN. Therefore, the currently best constraints on the one-halo-term by clustering measurements comes from a study with cross-correlations between nearby ($z \lesssim 0.5$) AGN and luminous-red-galaxies (LRGs), which is a strongly clustered galaxy class (e.g. Krumpe et al. 2010b, 2012; Miyaji et al. 2011). Other constraints come from the direct measurement of the mean halo occupation distribution (HOD) of galaxies and AGN in the COSMOS field (e.g. Allevato et al. 2014). In general, the constraints are statistically not very significant. They nevertheless show that AGN reside in DMHs with a mass of at least $\sim 10^{13} h^{-1} M_{\odot}$. At this mass threshold, AGN reside preferentially at the center of DMH, while for higher DMH masses AGN are also found as a satellite source in the DMH. This could be an indication that a satellite AGN has a different triggering mechanism than a central AGN. The studies also indicate that within one DMH the number of AGN decreases with the DMH mass in comparison to the total number of galaxies in the same DMH.

In order to increase the significance for measuring the one-halo-term, significantly larger AGN samples are needed. However, as we will show in Chapter 4 clustering studies from brightness fluctuations of the unresolved CXB for angular scales of $\lesssim 2'$ might also offer a good opportunity to study the one-halo-term of AGN.

1.4 X-ray telescopes

X-ray astronomy provides us an unique window to the Universe.

X-ray telescopes enable us to see the hottest gas of the Universe with temperatures of the order of tens of million degrees (keV-regime). Such temperatures are reached via shock-heating of various in- and outflows of matter. The brightest X-ray emission comes from the intracluster medium (ICM), which is the hot gas within DMHs of galaxy clusters & groups. They represent the largest and most massive structures of the Universe, where most of the baryonic matter is locked up. X-ray telescopes give us an exceptional way to study these structures and their containing matter by mapping the emission of hot gas in the extragalactic sky (e.g. with the eROSITA all-sky survey, Sect. 1.5).

The other major X-ray source are accreting compact objects, where gravitational energy is very efficiently transformed into radiation. The bulk of the emission is being emitted in the UV and X-ray regime. The most luminous emission from this process is produced by the accreting SMBH in AGN (Sect. 1.2). X-ray telescopes enable us to study how SMBHs grow and how they influence their environment over

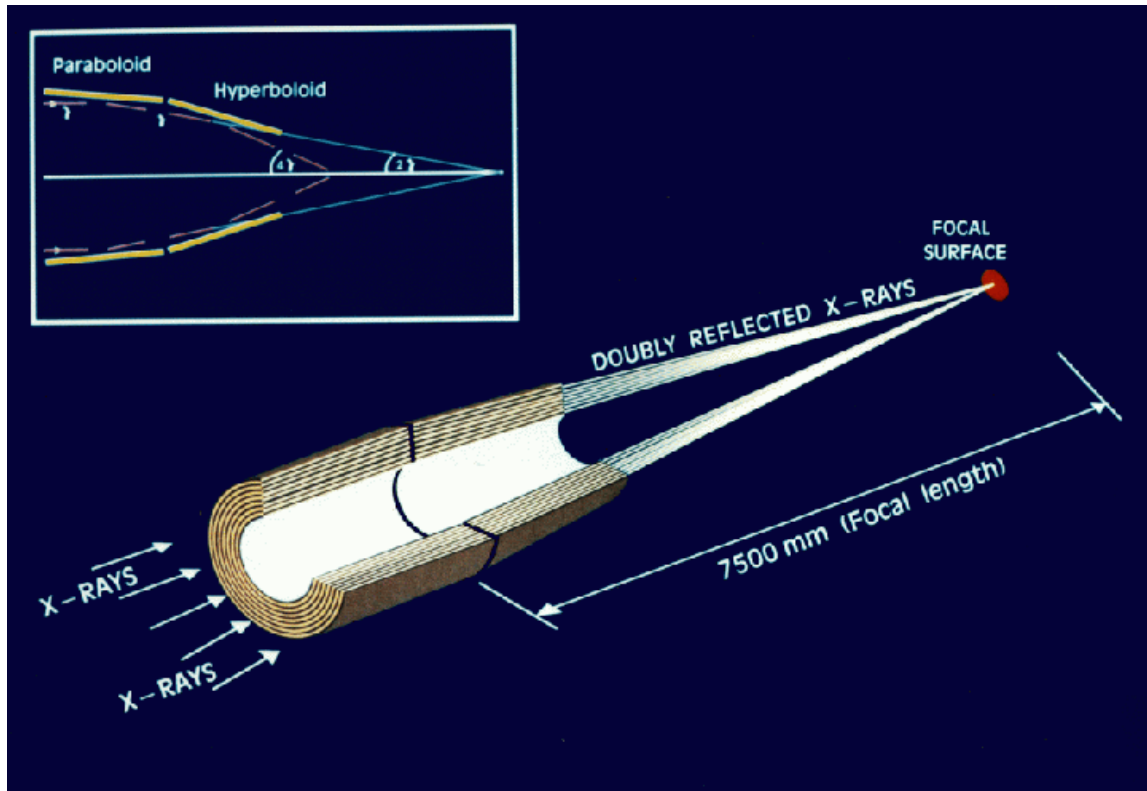


Figure 1.3: Focusing mechanism of an X-ray telescopes on the example of XMM-*Newton* using grazing incidence optics with a combination of a hyperbolic and parabolic mirrors to make use of the very small reflection angle of X-rays. (from “XMM-Newton Users Handbook”, ESA, XMM-Newton SOC, Fig. 2).

cosmic time.

The atmosphere of Earth absorbs X-rays completely. In consequence all X-ray telescopes must be space-based. The first orbital observatory specifically designed for X-ray observations was UHURU¹¹ in 1970, followed by many others, such as HEAO 1, Einstein, EXOSAT, and ROSAT (for a full list see e.g. Bradt et al. 1992). The current generation¹² of X-ray telescopes is lead by *Chandra* and XMM-*Newton* for the 0.5 – 10 keV regime, which is the energy-range of interest for the work presented in this Thesis.

The X-ray telescopes are different from optical telescopes, where the major difference comes from the focusing mechanism of the radiation. While an optical telescope can use one parabolic mirror perpendicular to the incoming light (*normal* incidence optics), X-ray telescopes have to use *grazing* incidence optics. Fig. 1.3 shows an example of such optics, which is composed of a system of nested, gold-coated, parabolic and hyperbolic mirrors to deflect the incoming X-rays by a small angle ($< 2^\circ$). A large number of nested mirrors helps to increase the effective area of the instrument.

The effective area is one of the main characteristics of an X-ray telescopes and is mainly given by the number and diameter of the mirror-shells. XMM-*Newton* with its 58 shells, compared to *Chandra*'s four shells, has currently the largest effective area of X-ray telescopes in the 0.5 – 10 keV regime.

The angular resolution is another main characteristic of an X-ray telescope. It is given by the surface smoothness and shape of the mirror and by the alignment of the mirror shells within the telescope. The better these properties, the more compact the point-spread-function (PSF), which describes the image of a single point source as obtained by the telescope optics. Since *Chandra* has the mirrors with the highest quality, it also has the smallest PSF of all X-ray telescopes with only 0.5'' (on-axis). In comparison, XMM-*Newton* and eROSITA have an on-axis PSF of $\sim 15''$.

For our analysis in Chapter 4, we use the XBOOTES survey, which was conducted by *Chandra*. More specifically, the survey was performed with the ACIS-I¹³ chips, which are an array of 2×2 of four 1024×1024 pixel CCDs. The high angular resolution of *Chandra* is one of the main reason why we chose to use this survey.

An interesting advantages of X-ray observations in compare to lower wavelength bands is that we can obtain the energy information for each detected X-ray photon with a resolution of ~ 100 eV (e.g. *Chandra* ACIS-I CCD¹³) without using any grating optics. This facilitates the analysis of X-ray observations. We also take advantage of the spectral information in our analysis with the XBOOTES survey in

¹¹<http://heasarc.gsfc.nasa.gov/docs/uhuru/uhuru.html>

¹²<http://heasarc.gsfc.nasa.gov/docs/heasarc/missions/active.html>

¹³<http://cxc.harvard.edu/proposer/POG/html/chap6.html>

Chapter 4. When using grating optics, one can improve the energy-resolution even further (e.g. *Chandra* HETG instrument¹⁴).

1.5 The eROSITA all-sky survey

The eROSITA¹⁵ all-sky survey (eRASS) will be the successor of RASS, the last all-sky survey conducted with ROSAT 25 years ago (see Fig. 1.1). The entire eROSITA mission is described comprehensively in the eROSITA Science Book (Merloni et al. 2012). Here, we summarize the most important aspects of the mission.

The eROSITA (extended ROentgen Survey with an Imaging Telescope Array) telescope will be the main instrument aboard the Russian Spektrum-Roentgen-Gamma satellite¹⁶, which is scheduled for launch in 2016. The eROSITA telescope has a similar effective area and PSF-size (on-axis) as XMM-*Newton* (Sect. 1.4) but its field-of-view (FOV) is with $\sim 1^\circ$ more than three times larger. It will scan the sky continuously for four years in the energy range of 0.5 – 10.0 keV. In the soft band (0.5 – 2.0 keV) eRASS will be about 30 times better in sensitivity than RASS and eRASS will perform the first all-sky survey ever in the hard band (2.0 – 10.0 keV).

The main science driver of eRASS is to study the accelerating expansion and the large-scales structure of the Universe, and the accretion history of SMBHs. This will be accomplished by creating the largest sample of galaxy clusters and AGN to date.

Galaxy clusters are a great probe of the large-scales structure and *dark energy* (DE) (as well as BAOs, Sect. 1.6). They are a long-known probe of *dark matter* (DM) (e.g. Zwicky 1933) and they gave us one of the first indications for the existence of DE (e.g. Peebles et al. 1989). Thanks to galaxy clusters, we are able to trace the most massive and largest structures in the Universe ($\gtrsim 10^{14} h^{-1} M_\odot$, $\gtrsim 1$ Gpc). Galaxy clusters are best detected in X-rays due their hot ICM with temperatures in the keV-regime ($\sim 10^7$ K, also see Sect. 1.4). Taking the X-ray flux and the well-calibrated scaling relations between the ICM temperature and the DMH mass (e.g. Vikhlinin et al. 2006; Sun et al. 2009), we can use the spatial density of galaxy clusters to determine the halo mass function (number density of DMHs as a function of mass) and compare it with predictions of different cosmological models. A comprehensive explanation of using galaxy clusters as a cosmological probe is given by Allen et al. (2011) and Weinberg et al. (2013). eRASS will detect almost all massive galaxy clusters in the X-ray sky, thus creating the largest X-ray detected galaxy cluster

¹⁴<http://cxc.harvard.edu/proposer/POG/html/chap8.html>

¹⁵<http://www.mpe.mpg.de/eROSITA>

¹⁶<http://hea.iki.rssi.ru/SRG>

sample with $\sim 10^5$ objects. It will improve the cosmological constraints by up to two orders of magnitude in comparison to the constraints of all previous X-ray detected galaxy cluster samples (see the eROSITA Science Book for more details).

eRASS will detect about $\sim 3 \times 10^6$ AGN, which are the subject of this Thesis in Chapter 2 and Chapter 3. In particular, in Chapter 2 we also give some more technical details about the eROSITA telescope and the strategy of its all-sky survey.

1.6 Baryon acoustic oscillations

With baryon acoustic oscillations (BAOs) we can study the accelerating expansion of the Universe as a function of redshift and therefore test different cosmological models, including DE. BAOs occur in the clustering signal of galaxies due a small DM over-density at a specific co-moving distance from the center of a DMH. This distance is called the *acoustic scale*. The size of this acoustic scale can be derived from the cosmological model and calibrated with the results from the power spectrum of the cosmic microwave background (CMB) temperature fluctuations (Planck Collaboration et al. 2015). This makes BAOs a standard ruler, with which we measure an absolute scale and compare it with predictions from theory.

This acoustic scale imprinted in the DM distribution in the early Universe transformed into a preferred clustering scale of galaxies at later times. This features manifest itself as an amplitude oscillations in the power spectrum of the galaxy distribution (top of Fig. 1.4). It also appears as an hump in the corresponding correlation function (bottom of Fig. 1.4). The BAOs in a pure baryonic universe were predicted to have a strong amplitude in the galaxy power spectrum (Sunyaev & Zeldovich 1970; Peebles & Yu 1970). However, with the discovery of DM it became clear that the signal must be much weaker. Therefore, only the breadth and depth of recent galaxy surveys, such as 2dFGRS (Fig. 1.2) and SDSS, allowed us to convincingly detect BAOs in the observed galaxy distribution (e.g. Fig. 1.4, Cole et al. 2005; Eisenstein et al. 2005; Hütsi 2006; Tegmark et al. 2006). Since then many larger and deeper galaxy surveys have followed (see Fig. 1.5) and BAOs have been measured up to $z \sim 1$ with galaxy surveys (e.g. Anderson et al. 2012). There are several dedicated BAO surveys planned or already in progress, such as eBOSS, BigBOSS, or HETDEX (also see Table 3.1), to improve the current BAO measurements and to expand to higher redshift ($z > 1$). A comprehensive description of BAOs can be found in the review of Bassett & Hlozek (2010) and also in Weinberg et al. (2013, Ch. 4).

Given the properties of BAOs, they became one of the major probes to study the accelerating expansion of the Universe, which is the most astonishing result of cosmology since the discovery of DM. Besides BAOs, Type Ia supernovae (SNe

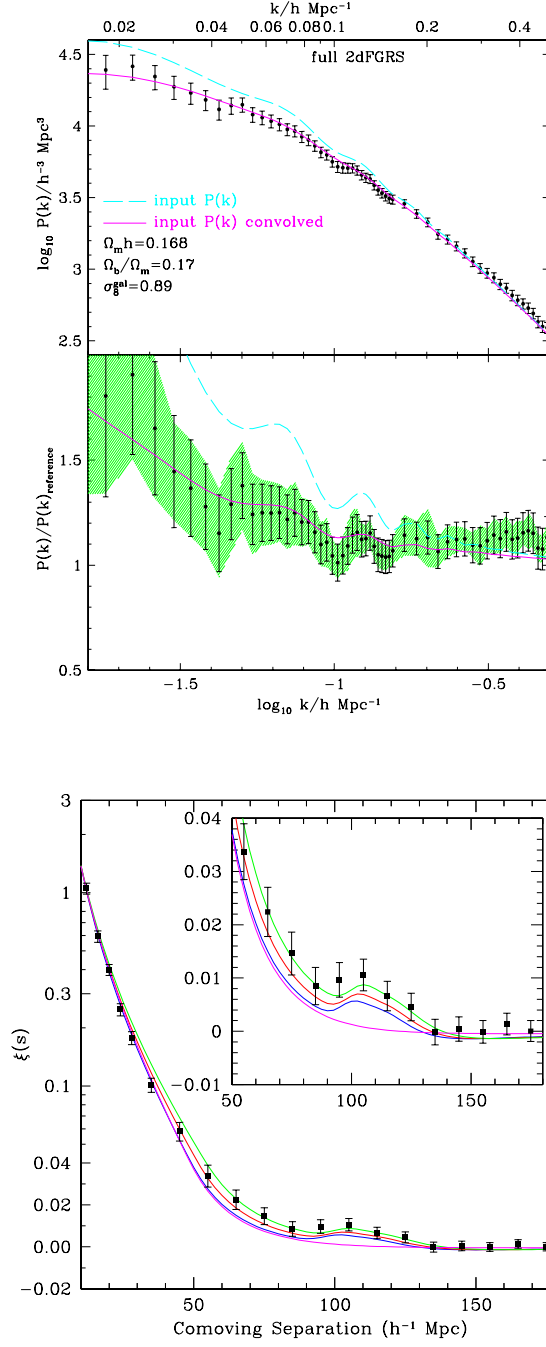


Figure 1.4: First convincing detection of BAOs in the clustering signal of galaxies. *Top:* The power spectrum of 2dFGRS galaxies (Cole et al. 2005, Fig. 12). The BAOs are visible as an amplitude oscillation in the power spectrum. The lower panel shows the measured power spectrum (including BAOs) divided by a theoretical power spectrum without BAOs (black crosses). *Bottom:* The correlation function of SDSS galaxies (Eisenstein et al. 2005, Fig. 2) with different cosmological models. The BAOs are visible as the hump around $110 h^{-1} \text{ Mpc}$.

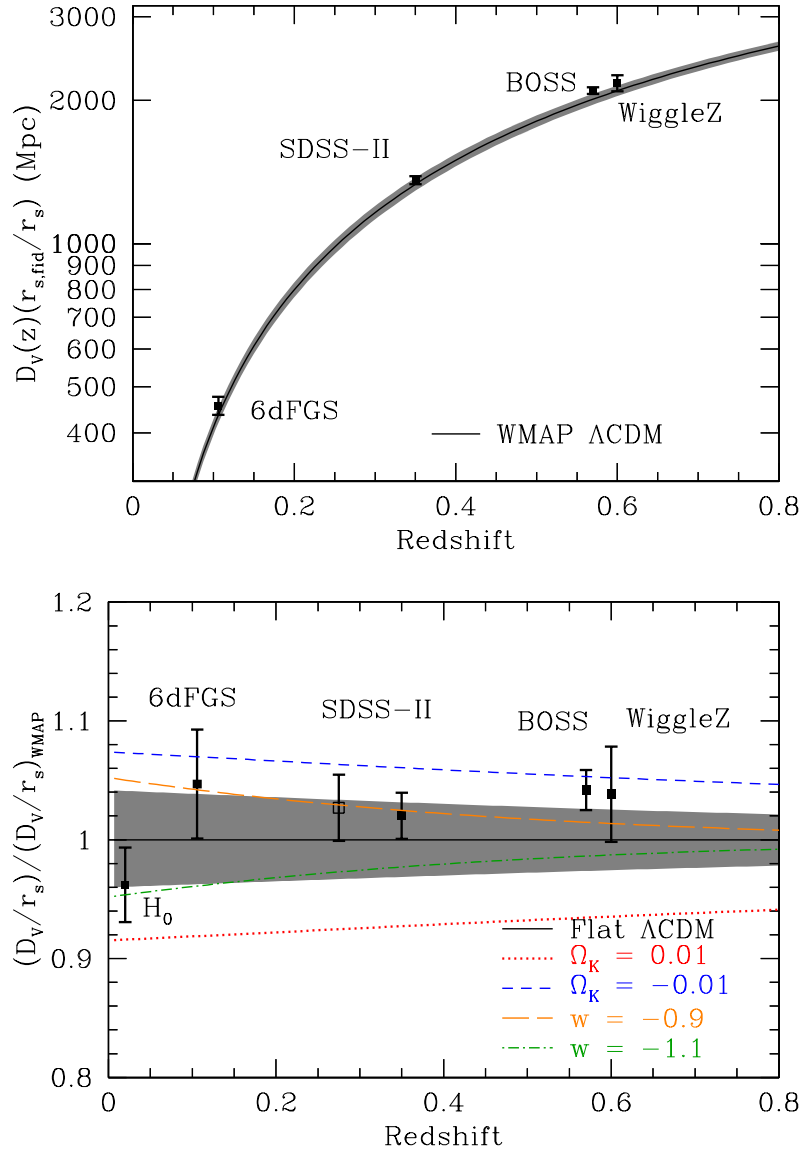


Figure 1.5: Comparison of recent BAO distance-redshift measurements from different galaxy surveys with cosmological models (Weinberg et al. 2013, Fig. 8). The quantity $D_V(z)$ combines the three-dimensional measurement of the acoustic scale from the line-of-sight (redshift space) and the transverse scale (Eisenstein et al. 2005, Eq. 2).

Ia), CMB, and galaxy clusters (see Sect. 1.5) are also major cosmological probes. Combining the results from these different probes reduces significantly degeneracies between cosmological model parameters and gives the currently best observational constraints on the properties of DE. One of the current results is that the density of DE appears to be constant over cosmic time (e.g. Frieman et al. 2008; Weinberg et al. 2013).

1.7 Overview of the Thesis

The Thesis is divided into two parts. In the first part, we focus on the prospects of using the eRASS-AGN sample for LSS studies. In Chapter 2, we study extensively the statistical properties of this sample. In Chapter 3, we investigate the potential of using the eRASS-AGN sample for LSS studies by concentrating on the clustering strength of AGN and the BAO measurement with AGN. The studies of these Chapters have been already published in Kolodzig et al. (2013b), Kolodzig et al. (2013a), and Hütsi et al. (2014). In the second part, we focus on LSS studies with the unresolved CXB. In Chapter 4, we study the surface brightness fluctuations of the unresolved CXB of the X-ray survey XBOOTES, the currently largest contiguous *Chandra* survey. In the final Chapter, we summarize the conclusions of the entire Thesis.

For the Thesis, we assume a flat Λ CDM cosmology with the following parameters: $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ($h = 0.70$), $\Omega_m = 0.30$ ($\Omega_\Lambda = 0.70$), $\Omega_b = 0.05$, $\sigma_8 = 0.8$. The values for H_0 and Ω_m were chosen to match the values assumed in the XLF studies, which we use in our calculations (e.g. Sect. 2.4), and Ω_b and σ_8 are derived from the CMB study of WMAP¹⁷ (Komatsu et al. 2011). We note that the results of this work are not very sensitive to the exact values of the cosmological parameters and if we would use the recently published, more precise cosmological parameters of the CMB study by PLANCK¹⁸ (Planck Collaboration et al. 2015) our results would not change.

We use the decimal logarithm throughout, if not stated otherwise.

¹⁷<http://map.gsfc.nasa.gov/>

¹⁸<http://www.cosmos.esa.int/web/planck>

Part I

Large-scale structure studies with AGN - prospects for eRosita

Chapter 2

Statistical properties of AGN in the eROSITA all-sky survey

2.1 Introduction

Large samples of X-ray detected active galactic nuclei (AGN) combined with follow-up optical data for identifying objects and their redshift determination are fundamental for understanding AGN evolution and the growth of supermassive black holes (SMBHs) with cosmic time. These samples are constructed in various extragalactic X-ray surveys spanning from wide-shallow to narrow-deep surveys. Many of these have been conducted in the past decade with the *Chandra* and XMM-*Newton* X-ray observatories, which were instrumental in understanding the cosmic X-ray background and evolution of AGN at low- and high redshifts (Brandt & Hasinger 2005). While *Chandra* and XMM-*Newton* have now surveyed several hundreds of square degrees (e.g. $\sim 360 \text{ deg}^2$ of the 2XMM-catalog, Watson et al. 2009), the sensitivity of the archival observations is far from being homogeneous. Moreover, the sky area covered by largest contiguous surveys did not exceed several (e.g. XMM-COSMOS: Hasinger et al. 2007; XBootes: Murray et al. 2005) to several tens of square degrees (e.g. XMM-LSS: Chiappetti et al. 2012; XWAS: Esquej et al. 2013; XXL: Pierre et al. 2011), which in particular limited our knowledge of the evolution and clustering properties of the most luminous AGN severely. The most recent *all-sky survey* (Voges et al. 1999) in the X-ray band was performed by ROSAT¹ (Truemper 1993) more than two decades ago, creating an increasing demand for an all-sky survey to be conducted by the new generation of X-ray telescopes.

¹<http://www2011.mpe.mpg.de/xray/wave/rosat/>

Table 2.1: Predicted background count rates.

Energy band [keV]	0.5 – 2.0	2 – 10
Particle	0.3	2.6
Galactic	1.8	0.0
Extragalactic	1.9	0.5
Total	4.0	3.1

The count rates are given in units of 10^{-4} cts s $^{-1}$ per PSF HEW for seven telescopes. The extragalactic component accounts for unresolved sources only and at the average four-year survey sensitivity (Table 2.2).

The eROSITA² telescope (extended ROentgen Survey with an Imaging Telescope Array) will be able to satisfy this demand. It is the main instrument aboard the Russian Spektrum-Roentgen-Gamma satellite³, which is scheduled for launch in 2016. Its main science goals are cosmological studies of clusters of galaxies and AGN, with the aim of constraining the nature of dark matter and dark energy. To achieve these goals, eROSITA will perform an all-sky survey (eRASS) during the first four years of its operation, followed by a phase of pointed observations. The main mission parameters and the telescope are described in the eROSITA Science Book (Merloni et al. 2012, hereafter SB).

In this study we explore the main statistical properties of the AGN sample that is expected to be detected in the course of the eRASS, including its luminosity- and redshift distributions. This will help to understand the capabilities of the eROSITA mission and, potentially, to tune the survey strategy and its parameters.

The calculations in this Chapter are performed for two energy bands – soft (0.5 – 2.0 keV) and hard (2.0 – 10.0 keV). In computing count rates we used the most recent response matrix of eROSITA, `erosita_iv_7telfov_ff.rsp`⁴. As is appropriate for the all-sky survey data analysis, this response matrix is averaged over the field-of-view and scaled to 7 telescopes. In this work, we assume that the data from the entire survey of eROSITA is available for analysis.

²<http://www.mpe.mpg.de/eROSITA>

³<http://hea.iki.rssi.ru/SRG>

⁴<http://www2011.mpe.mpg.de/erosita/response/>

Table 2.2: Characteristic average parameters of the eROSITA all-sky survey.

Survey duration	4.0 years				0.5 years	
	Extragalactic sky		Ecliptic poles		Extragalactic sky	
Sky region	34 100 ($ b > 10^\circ$)		90		34 100 ($ b > 10^\circ$)	
Solid angle [deg ²]	34 100 ($ b > 10^\circ$)		90		34 100 ($ b > 10^\circ$)	
Exposure time [sec]	2 000		20 000		260	
Energy band [keV]	0.5 – 2.0	2 – 10	0.5 – 2.0	2 – 10	0.5 – 2.0	2 – 10
Resolved extrag. CXB [%]	31	6	53	17	12	$\lesssim 1$
Background counts [cts/PSF]	0.8	0.6	6.7	6.1	0.1	0.1
Source counts [cts/PSF]	7.6	6.8	16.5	15.6	4.4	3.9
Sensitivity $\langle S_{\text{lim}} \rangle$ [10^{-14} erg s ⁻¹ cm ⁻²]	1.1	18	0.23	4.2	4.8	80
Source density [deg ⁻²]	84	3.7	450	37	10.0	0.4
Number of sources [$\times 10^3$]	2 900	130	41	3.4	340	13

2.2 Sensitivity

The point-source detection sensitivity of eROSITA in the all-sky survey was discussed in detail by Prokopenko & Gilfanov (2009). Since then, the spacecraft orbit has been changed to the L2 orbit and detailed calculations of the instrumental background became available. We therefore update the calculations of Prokopenko & Gilfanov (2009) below. We first compute the sky average values (Sect. 2.2.1-2.2.3) and then calculate a more realistic sensitivity map (Sect. 2.2.4), which is used for our eRASS AGN calculations.

2.2.1 Instrumental and cosmic background

The eROSITA background is dominated by the photon background below ~ 2 keV and by the particle component above this energy. In the subsequent calculations we used the field-of-view-averaged angular resolution (PSF⁵ HEW⁶) of $30''$ diameter for the soft band and $40''$ diameter in the hard band (Friedrich et al., priv. comm.). The corresponding PSF HEW areas are ≈ 707 and $\approx 1\,257$ arcsec², respectively. For the eROSITA focal length, 1 mm on the detector corresponds to ≈ 128.8 arcsec.

The instrumental non X-ray (particle) background spectrum is nearly flat in the counts space, with a normalization of $\approx 6.1 \times 10^{-3}$ cts s⁻¹ cm⁻² keV⁻¹ (Perinati

⁵PSF - Point-spread-function

⁶HEW - Half-energy width, alias half-power diameter (HPD)

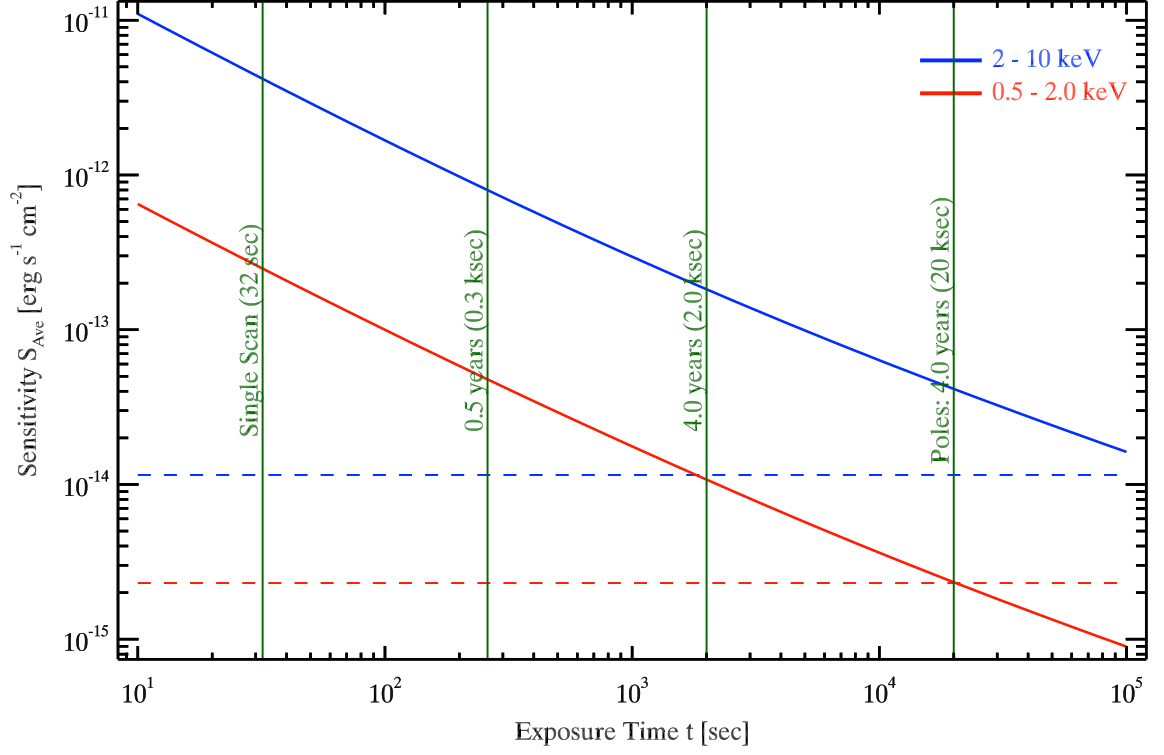


Figure 2.1: Average point-source detection sensitivity of the eROSITA telescope as a function of the exposure time in the soft (red lines, 0.5 – 2.0 keV) and hard (blue lines, 2.0 – 10.0 keV) energy bands for the extragalactic sky. The horizontal dashed lines are the confusion limits for one source per 40 telescope beams (PSF HEW). The vertical line on the left shows the exposure time for a single scan at 100 % observing efficiency, the other vertical lines indicate the average exposure times for different survey durations and at the ecliptic poles (for the soft-band confusion limit) at 80 % observing efficiency.

et al. 2012). Within the PSF HEW it produces count rates of $\approx 2.7 \times 10^{-5}$ and $\approx 2.6 \times 10^{-4}$ cts s $^{-1}$ for the soft and hard band for seven telescopes. Solar-activity-induced background events are not taken into account in this calculation, they are roughly accounted for via the observing efficiency introduced in the survey exposure time calculations. These numbers are the result of purely theoretical calculations and there are no direct measurements of the real background of an X-ray detector in the L2 orbit. Therefore the above numbers may have to be revised after the eROSITA launch.

The X-ray photon background (CXB) has two components (Lumb et al. 2002): (1) the truly diffuse Galactic background of local ionized ISM⁷ with a soft thermal spectrum, and (2) the hard power-law extragalactic CXB component.

To estimate the contribution of the ionized ISM emission we used the spectral fits from Lumb et al. (2002, Table 3) and obtained a count rate of $\approx 1.8 \times 10^{-4}$ cts s $^{-1}$ within the PSF HEW for the soft band, assuming the same Galactic absorption ($N_{\text{H}} = 1.7 \times 10^{20}$ cm $^{-2}$) and solar abundance (Anders & Grevesse 1989) as Lumb et al. (2002). Its contribution to the hard band can be neglected. As a caveat we note that the Lumb et al. (2002) analysis was based on the XMM-*Newton* observations of several blank fields located at high Galactic latitudes. Therefore these numbers should be considered as approximate, because they do not account for inhomogeneity the of the Galactic background radiation.

For the extragalactic CXB component we assumed a power-law spectrum with a photon index of $\Gamma = 1.42$ (Lumb et al. 2002, Table 3). The power-law was normalized to the extragalactic CXB flux of 7.53×10^{-12} and 2.02×10^{-11} erg s $^{-1}$ cm $^{-2}$ deg $^{-2}$ for the soft and hard band (Moretti et al. 2003)⁸. We furthermore assumed Galactic absorption of $N_{\text{H}} = 6 \times 10^{20}$ cm $^{-2}$ corresponding to the arithmetic mean of the N_{H} -map of Kalberla et al. (2005) for the extragalactic sky ($|b| > 10^\circ$). With these parameters, the average count rate of extragalactic CXB within the PSF HEW (seven telescopes) is $\approx 2.8 \times 10^{-4}$ and $\approx 5.8 \times 10^{-5}$ cts s $^{-1}$ in the soft and hard band, respectively.

In computing the contribution of extragalactic sources, one needs to take into account that a fraction of the background AGN will be resolved by eROSITA. Therefore, these sources will not contribute to the unresolved image background, which will affectg the point-source detection sensitivity. This effect reduces of the extragalactic

⁷ISM - Interstellar medium

⁸Strictly speaking, these flux values correspond to a slightly steeper slope (≈ 1.45) of the CXB spectrum than the conventional value of 1.42. This discrepancy reflects the uncertainty in the absolute CXB flux determinations. We nevertheless used them for consistency with the resolved fraction calculations below.

background count rate. At the average four year survey sensitivity of eROSITA (see Sect. 2.2.2) the resolved extragalactic CXB fraction⁹ achieves $\approx 31\%$ in the soft and $\approx 6\%$ in the hard band. The fractions were calculated using the number counts of Georgakakis et al. (2008, hereafter G08) and the extragalactic CXB flux of Moretti et al. (2003). Thus, the final values of the average count rate of the unresolved CXB emission within the PSF HEW is $\approx 1.9 \times 10^{-4}$ and $\approx 5.4 \times 10^{-5}$ cts s⁻¹ in the soft and hard band, respectively.

To examine the background model, we used ROSAT all-sky maps of diffuse X-ray emission (Snowden et al. 1997). We combined PSPC maps¹⁰ in the energy bands from *R4* to *R7* (see Snowden et al. 1997, Table 1) to cover the 0.44–2.04 keV range, which presents a reasonable approximation of the eROSITA soft band. The median count rate on the combined map is ≈ 0.87 PSPC cts s⁻¹ deg⁻² for the extragalactic sky ($|b| > 10^\circ$). A convolution of our X-ray background model with the ROSAT PSPC response matrix (`pspcb_gain2.256.rsp`¹¹) gives a 0.44–2.04 keV count rate of ≈ 0.99 cts s⁻¹ deg⁻², which is sufficiently close to the median value in the ROSAT map. ROSAT maps show moderate brightness variations in which 90 % of the extragalactic sky count rate are between ≈ 0.34 and ≈ 2.17 PSPC cts s⁻¹ deg⁻². Variations of the background count rate in this range amount to $\sim 30\%$ variations in the sensitivity.

The contributions of different background components are summarized in Table 2.1. They are consistent within $\approx 10\%$ with the numbers in the SB. The difference in the soft band comes from the slight difference in the normalization of the extragalactic component. The change in the hard band appears because we used the results of the updated particle background calculations of Perinati et al. (2012) instead of those of Tenzer et al. (2010) which were used in the SB.

2.2.2 Average exposure and sensitivity

With the average background count rates we computed the point-source detection sensitivity of eROSITA as a function of the exposure time, which is shown in Fig. 2.1. In this computation we assumed a Poissonian distribution of counts and demanded no more than 200 false point-source detections for the entire sky. This corresponds to one false detection in ≈ 250 fields of view (≈ 210 deg²). For a Gaussian distribution, this false-detection rate is equivalent to an $\approx 5.0\sigma$ confidence level in one trial. In converting the count rates into flux we assumed an absorbed power-law spectrum

⁹ Note that Fig. 5 in Moretti et al. (2003) gives resolved fractions of $\sim 50\%$ and $\sim 10\%$, respectively. This difference also reflects the uncertainties of the CXB measurements.

¹⁰ <http://www.xray.mpe.mpg.de/rosat/survey/sxrb/12/ass.html>

¹¹ <ftp://legacy.gsfc.nasa.gov/caldb/data/rosat/pspc/cpf/matrices/>

with a photon index $\Gamma = 1.9$ and a sky-averaged Galactic absorption of $N_{\text{H}} = 6 \times 10^{20} \text{ cm}^{-2}$. We also took into account that only half of the source counts are contained within the PSF HEW.

The sky-averaged exposure time of the survey is

$$t_{\text{exp}} \approx 2.0 \left(\frac{t_{\text{survey}}}{4 \text{ years}} \right) \left(\frac{f_{\text{eff}}}{80 \%} \right) \left(\frac{\text{FOV}}{0.833 \text{ deg}^2} \right) \text{ ksec} , \quad (2.1)$$

where t_{survey} is the survey duration and f_{eff} is the observing efficiency, whose expected value is $f_{\text{eff}} \approx 80 \%$, and the eROSITA field-of-view (FOV) is 1.03° in diameter (see the SB). The average numbers of background counts to be accumulated within the PSF HEW in the course of the four-years survey (average exposure time of ≈ 2.0 ksec per point) are ≈ 0.8 and ≈ 0.6 in the soft and hard band. For these numbers and for the chosen confidence level, the source detection thresholds are ≈ 8 and ≈ 7 source counts within the PSF HEW. The corresponding point-source detection sensitivities are $\langle S_{\text{lim}} \rangle \approx 1.1 \times 10^{-14}$ and $\approx 1.8 \times 10^{-13} \text{ erg s}^{-1} \text{ cm}^{-2}$ in the soft and hard band.

After the first half year of the survey, eROSITA will have scanned the whole sky once. At the averaged exposure time of ≈ 260 sec, there will be ≈ 0.1 background counts per PSF HEW in each energy band and the point-source detection threshold will be ≈ 4 counts. The point-source detection limits for the half-year survey are $\approx 4.8 \times 10^{-14}$ and $\approx 8.0 \times 10^{-13} \text{ erg s}^{-1} \text{ cm}^{-2}$.

The main characteristics of the full survey and its first half year are summarized in Table 2.2. The numbers are generally consistent with the SB. The small differences are related to the differences in the background estimates and the larger PSF size used here for the hard band. We also iteratively calculated the resolved fraction of the extragalactic CXB for sensitivities.

Thus, the eRASS will have on average an ~ 30 times better sensitivity in the soft band than the previous all-sky survey in this band conducted by ROSAT (Voges et al. 1999). On the other hand, its sensitivity is between one to four orders of magnitude lower than that of the deep but much more narrow (some of them pencil-beam) extragalactic X-ray surveys conducted by *Chandra* and *XMM-Newton*, such as CDFs, COSMOS, Bootes, Lockman Hole, and ChaMP (see Brandt & Hasinger 2005, for a review).

2.2.3 Confusion limit

For the purpose of this study we assumed that the source confusion becomes important at a source density of one sources per 40 telescope beams (= PSF HEW), which for the angular resolution of eROSITA corresponds to a source density of ≈ 460 and

≈ 260 sources deg^{-2} in the soft and hard band. With the average of the extragalactic $\log N - \log S$ curves of G08 and Kim et al. (2007, Table 3, ChaMP+CDFs, hereafter K07), the corresponding flux levels are $\approx 2.3 \times 10^{-15}$ and $\approx 1.2 \times 10^{-14}$ $\text{erg s}^{-1} \text{cm}^{-2}$. In the soft band the confusion limit is achieved at an exposure time of ≈ 20 ksec. In the hard band, source confusion, achieved at an exposure time of ≈ 190 ksec, is not a problem for eRASS.

2.2.4 Sensitivity map

Owing to the properties of the scan pattern, the two ecliptic poles of eROSITA will have a significantly higher exposure time than the sky on average. This leads to a higher sensitivity at the ecliptic poles. The scan strategy of eROSITA is still under discussion and different scenarios are still possible, depending on whether the satellite rotation axis is continuously pointing at the Sun or moves around it with a slight offset. In the latter scenario, the ecliptic pole regions will occupy a larger solid angle and will be less overexposed. We consider here the extreme case of the continuous Sun-pointing of the scan axis.

Using the exposure map of the four-year survey (Robrade, priv. comm.), we defined the two sky regions at the ecliptic poles of eROSITA, where the exposure time (reduced by the observing efficiency) exceeds the confusion limit of 20.0 ksec. These two pole regions combined cover a solid angle of ≈ 90 deg^2 . The point-source detection sensitivity in the soft band in the pole regions is defined by the confusion limit and is approximately 2.3×10^{-15} $\text{erg s}^{-1} \text{cm}^{-2}$, taking into account $\approx 50\%$ resolved CXB fraction. The survey characteristics for the pole regions are summarized in Table 2.2. In the hard band, the confusion limit is reached at a much longer exposure time of ≈ 190 ksec and is not relevant for the all-sky survey. The actual hard-band sensitivity in the pole regions is determined by the particular scan pattern. For reference, we list in the Table 2.2 the sensitivity that can be achieved in the hard band assuming the 20.0 ksec exposure.

Outside the poles, the exposure time still varies significantly, with lowest value of ≈ 1.6 ksec achieved in the equatorial regions. These variations will lead to variations of the point-source detection sensitivity across the sky. To compute a realistic sensitivity map of the survey, we took into account variations of the Galactic absorption across the sky along with the exposure map. To this end we used the N_{H} -map of Kalberla et al. (2005). We excluded the Galactic plane and only considered the sky at Galactic latitudes $|b| > 10^\circ$ for the subsequent calculations. In computing the exposure map we assumed an observing efficiency of $f_{\text{eff}} = 80\%$ and set overexposed regions at the Galactic poles to 20.0 ksec. The solid angle of this extragalactic sky is

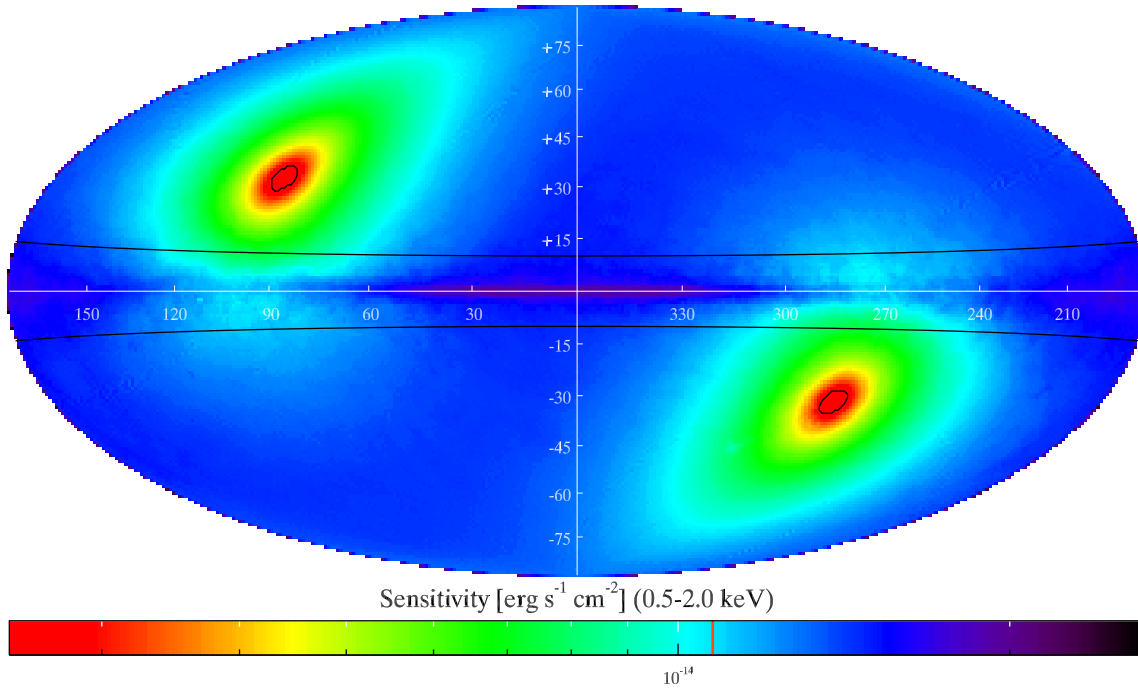


Figure 2.2: Four-year soft-band ($0.5 - 2.0$ keV) sensitivity map of eRASS in Galactic coordinates (l, b), based on the exposure time map of J. Robrade (priv. comm.) for a continuous Sun-pointing and based on the N_{H} -map of Kalberla et al. (2005). The two black horizontal curves enclose the Galactic plane ($|b| < 10^\circ$), which is excluded from our calculation, and the two regions encircled by black curves are our defined ecliptic poles, where the exposure time was set to 20.0 ksec. The red vertical line in the horizontal color bar shows the average sensitivity (from Table 2.2).

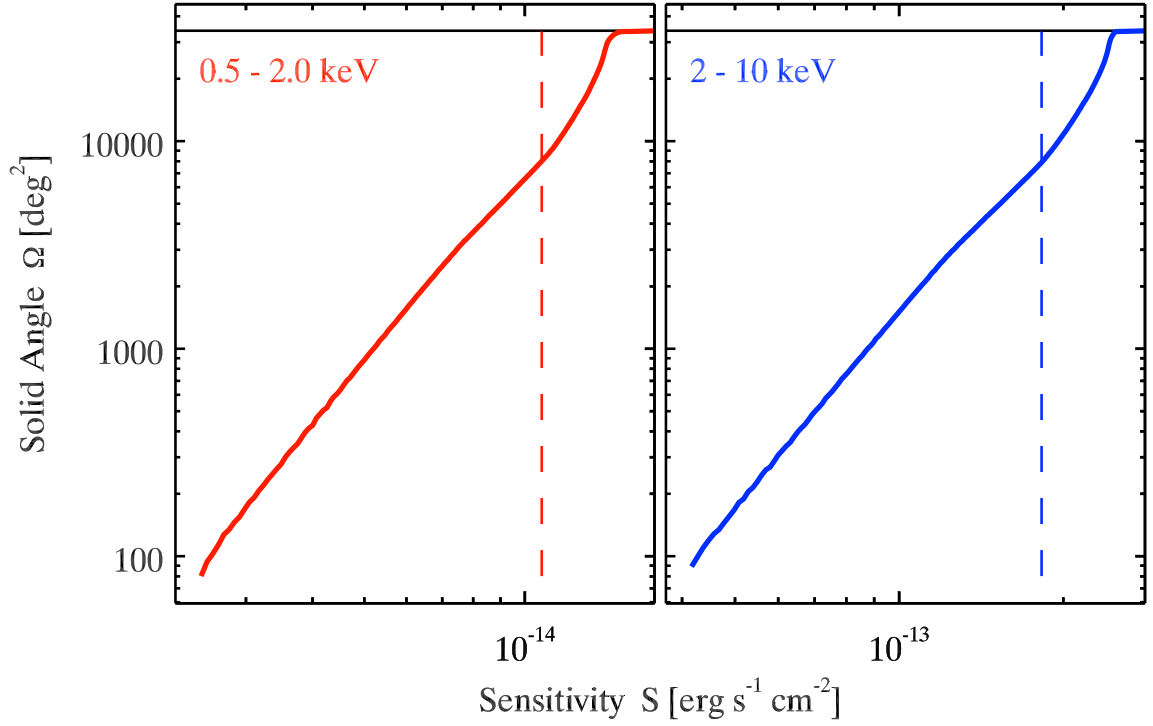


Figure 2.3: The sensitivity – solid angle curves for the soft (*left panel*) and hard (*right*) bands. The vertical dashed lines show the corresponding average sensitivities from the Table 2.2. The horizontal line shows the solid angle of the extragalactic sky.

$\Omega \approx 34\,100 \text{ deg}^2$, which corresponds to $\approx 83\%$ of the total sky. For the extragalactic sky the arithmetic mean of the exposure map is $\approx 2.1 \text{ ksec}$, which is close to the average exposure time computed from Eq. (2.1).

For background calculations, we assumed a constant count rate for the particle background. We assumed that the soft Galactic background is produced in the Local Bubble and therefore is not subject to Galactic absorption, whereas the contribution of the extragalactic CXB component was computed with the N_{H} -map taking into account. The resolved extragalactic CXB fraction was fixed at the sky-average value (see Sect. 2.2.1 and Table 2.2). We ignored intrinsic variations of the Galactic and extragalactic CXB, which are unrelated to Galactic absorption,. The amplitude of their variations across the sky can be inferred from the ROSAT diffuse sky maps, as described in Sect. 2.2.1. Background count-rate variations of such amplitude will result in sensitivity variations of $\sim 30\%$. A part of these variations is caused by the variations of the Galactic absorption and is included in our calculations through the N_{H} -map.

With these assumptions we computed sensitivity maps for the two bands; the one for the soft band is shown in Fig. 2.2. The sensitivity – solid angle dependences for both bands are shown in Fig. 2.3.

2.3 AGN number counts

To estimate the source densities and the total numbers of sources detected in different bands we used the source counts results of K07 and G08, displayed in Fig. 2.4. For K07, we used the best-fit parameters for the ChaMP+CDFs data from their Table 3 and converted the break flux (S_{b}) into the hard band from $2.0 - 8.0 \text{ keV}$ to $2.0 - 10.0 \text{ keV}$ assuming a power-law spectrum with a photon index of $\Gamma = 1.4$, as was used by K07. The best-fit parameters of K07 and G08 agree well (within 2σ). The difference in number counts between the two $\log N - \log S$ curves is mostly below 10% , the strongest deviation is about 20% in the flux range of interest (defined by the characteristic sensitivities, see Table 2.2). In the following calculations we used the average of the values given by the two $\log N - \log S$ curves.

With these curves and our sensitivity map in each energy band, we computed the number density map. The arithmetic mean of this map gives us an average number density of $\approx 81 \text{ deg}^{-2}$ and $\approx 3.8 \text{ deg}^{-2}$ in the soft and hard band. The total numbers of sources detected are $\approx 2.7 \times 10^6$ and $\approx 1.3 \times 10^5$ for the extragalactic sky. These values differ slightly from those in Table 2.2 because the latter were computed using the average sensitivities of the survey. They agree reasonably well with those in the SB. About 10% of these sources in both bands will be detected after the first half

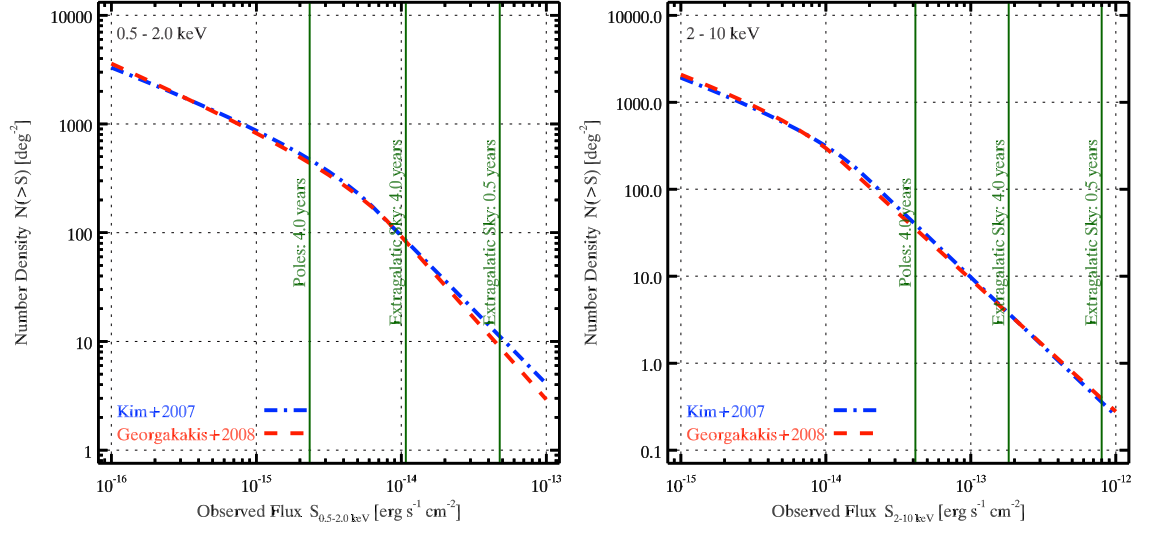


Figure 2.4: Cumulative number counts $N(>S)$ for observed X-ray point-sources in the soft (*left*) and hard band (*right*). The blue dashed-dotted curve is from Kim et al. (2007, Table 3, ChaMP+CDFs) and the red dashed curve from Georgakakis et al. (2008). The vertical solid lines indicate the point-source detection sensitivity for different survey durations assuming 80 % observing efficiency, and the sensitivity at the 20 ksec exposure time, corresponding to the confusion limit in the soft band (left-most line marked "Poles").

year of the survey.

Taking the corresponding sensitivities from Table 2.2, we computed a number density of $\approx 450 \text{ deg}^{-2}$ and $\approx 37 \text{ deg}^{-2}$ in the soft and hard band for the ecliptic poles. This translates into $\approx 41\,000$ and $\approx 3\,400$ detected extragalactic point-sources after four years.

About 10% of the brightest AGN in eRASS will be detected with at least ≈ 38 and ≈ 30 counts per PSF HEW (corresponding flux limits $\approx 5.4 \times 10^{-14}$ and $\approx 8.0 \times 10^{-13} \text{ erg s}^{-1} \text{ cm}^{-2}$) in the soft and hard band. The faintest 10% will have approximately 8 and 7 counts per PSF HEW ($\approx 1.2 \times 10^{-14}$ and $\approx 2.0 \times 10^{-13} \text{ erg s}^{-1} \text{ cm}^{-2}$).

We used the sensitivity map to estimate the numbers of AGN expected to be detected in the Galactic plane, $|b| < 10^\circ$, and obtained $\approx 4.6 \times 10^5$ and $\approx 2.2 \times 10^4$ sources. This is a somewhat lower number than predicted using the average source density on the extragalactic sky ($\approx 6.0 \times 10^5$ and $\approx 2.7 \times 10^4$) because the Galactic absorption in the Galactic plane is on average an order of magnitude higher than for the extragalactic sky. The higher Galactic X-ray background, not accounted for in these calculations, will additionally reduce the number of AGN at low Galactic latitudes, and high confusion with Galactic sources will make identifying them more difficult.

Finally, we note that AGN will be the most abundant source in eRASS. In addition, eRASS will detect about $\sim 10^5$ galaxy clusters (Predehl et al. 2010), $\sim 2 \times 10^4$ normal galaxies (Prokopenko & Gilfanov 2009), and $\sim 4 \times 10^5$ stars (see the SB).

2.4 X-ray luminosity function of AGN

With the knowledge of the X-ray luminosity function (XLF) of AGN,

$$\phi(L, z) = \frac{d\Phi(L, z)}{d\log L} \quad , \quad (2.2)$$

we are able to compute the redshift and luminosity distributions of the AGN expected to be detected in eRASS. The XLF describes the number of AGN (N) per unit co-moving volume (V) and logarithmic X-ray luminosity ($\log L$) as a function of X-ray luminosity L and redshift z . It is currently believed that the luminosity-dependent density evolution (LDDE) model describes the shape of the observed XLF best (Miyaji et al. 2000; Ueda et al. 2003; Hasinger et al. 2005; Silverman et al. 2008; Ebrero et al. 2009). For completeness, we summarize it below. The LDDE

model that parameterizes the AGN XLF has a double power-law

$$\phi(L, z) = K_0 \left[\left(\frac{L}{L_*} \right)^{\gamma_1} + \left(\frac{L}{L_*} \right)^{\gamma_2} \right]^{-1} e(L, z) \quad , \quad (2.3)$$

with the density evolution factor given by

$$e(L, z) = \begin{cases} (1+z)^{p_1} & z \leq z_c(L) \\ (1+z_c(L))^{p_1} \left(\frac{1+z}{1+z_c(L)} \right)^{p_2} & z > z_c(L) \end{cases} \quad , \quad (2.4)$$

where the cutoff redshift is given by

$$z_c(L) = \begin{cases} z_{c,0} \left(\frac{L}{L_\alpha} \right)^\alpha & L \leq L_\alpha \\ z_{c,0} & L > L_\alpha \end{cases} \quad , \quad (2.5)$$

This LDDE model has nine parameters. Miyaji et al. (2000) defined the density evolution factor (Eqs. 2.4 and 2.5) for the soft-band XLF in a slightly different way, but the concept remains the same. Hasinger et al. (2005) used the luminosity-dependent density evolution indices (p_1 and p_2)

$$p_1(L) = p_{1_{44}} + \beta_1 (\log L - 44.0) \quad (2.6)$$

$$p_2(L) = p_{2_{44}} + \beta_2 (\log L - 44.0) \quad , \quad (2.7)$$

and therefore the number of parameters increases with the two additional parameters (β_1 and β_2) to eleven.

As our default XLF models we used the LDDE model of Hasinger et al. (2005, Table 5, hereafter H05) for the soft band and of Aird et al. (2010, hereafter A10) for the hard band. For the hard-band XLF we used the best-fit model from A10, the "color preselected sample" (their Table 4), which is expected to provide a more accurate description of the XLF at higher redshifts. The parameters of the chosen XLF models are summarized in Table 2.3.

Based on the XMM-*Newton* data, Brusa et al. (2009) demonstrated that the soft-band XLF of H05 overpredicts the numbers of high-redshift objects, $z > 3$, detected in the COSMOS survey. These authors proposed to introduce an exponential redshift cutoff of the XLF for $z > 2.7$,

$$\phi = \phi_{\text{H05}}(z=2.7) \times 10^{0.43(2.7-z)}, \quad z > 2.7 \quad , \quad (2.8)$$

and showed that with this modification, the observed number counts of high-redshift AGN are reproduced much better. This result was also confirmed by Civano et al.

Table 2.3: Parameters of the LDDE model used to compute the luminosity and redshift distributions of the detected AGN.

Energy band [keV]	0.5 – 2.0	2 – 10
XLF	H05 (Table 5)	A10 (Table 4)
K_{44} / K_0 ^(a)	2.62 ± 0.16 ^(b)	8.32 ± 1.15
$\log L_*$ ^(c)	43.94 ± 0.11	44.42 ± 0.04
γ_1	0.87 ± 0.10	0.77 ± 0.01
γ_2	2.57 ± 0.16	2.80 ± 0.12
p_{144} / p_1	4.7 ± 0.3 ^(d)	4.64 ± 0.24
p_{244} / p_2	-1.5 ± 0.7 ^(d)	-1.69 ± 0.12
$z_{c,44} / z_{c,0}$	1.42 ± 0.11 ^(e)	1.27 ± 0.07
$\log L_\alpha$ ^(c)	44.67 (fixed)	44.70 ± 0.12
α	0.21 ± 0.04	0.11 ± 0.01
β_1	0.7 ± 0.3	–
β_2	0.6 ± 0.8	–

^(a) In units of 10^{-7} Mpc^{-3}

^(b) $K_0 = K_{44} [(10^{44.0}/L_*)^{\gamma_1} + (10^{44.0}/L_*)^{\gamma_2}] \approx 6.69$

^(c) erg s^{-1}

^(d) p_1 and p_2 are computed from Eqs. (2.6) and (2.7)

^(e) $z_{c,0} = z_{c,44} 10^{\alpha (\log L_\alpha - 44.0)} \approx 1.96$

(2011), who used additional *Chandra* data on the same field and analyzed an $\approx 50\%$ larger AGN sample than Brusa et al. (2009), and by Hiroi et al. (2012), who analyzed 30 high-redshift ($z > 3$) AGN in the Subaru/XMM-*Newton* Deep Survey field. Introducing the redshift cutoff results in an insignificant decrease of $\sim 1\%$ in the total number of AGN above the eRASS sensitivity limit. However, it has a strong effect on the numbers of high-redshift objects, which we discuss in Sect. 2.5.4. For our default XLF in the soft band, we included the high-redshift cutoff described by Eq. (2.8), but additionally show results without cutoff.

As a consistency check, we computed the $\log N - \log S$ distributions based on the chosen XLF models and compared them with the results of the source counts by K07 and G08. The $\log N - \log S$ can be computed by integrating the XLF over luminosity L and redshift z :

$$N(>S) = \int_0^{z_{\max}} \frac{dV(z)}{dz} \int_{\log L_{\min}(S,z)}^{\log L_{\max}} \phi(\log L, z) d\log L dz \quad . \quad (2.9)$$

Here, $\frac{dV(z)}{dz}$ [Mpc³ sr⁻¹] is the co-moving volume element per redshift and solid angle¹² and $L_{\min}(S, z) = 4\pi S d_L^2(z)$, where $d_L(z)$ is the luminosity distance (e.g. Hogg 1999). K-correction was applied, assuming a power-law spectra with the photon index $\Gamma = 1.9$ and no absorption. The same photon index was used to convert the XLFs to the energy bands used in this paper, if the former was determined for a different energy band. It is worth to mention that deep X-ray surveys do not show any evidence of a redshift-dependent photon index (Brandt & Hasinger 2005).

In Eq. (2.9) as well as in the calculations described in the next Sections, we integrated the XLF model in the luminosity range of $40 \leq \log L[\text{erg s}^{-1}] \leq 48$ and in the redshift range of $0 \leq z \leq 7$. Decrease of the L_{\min} in the luminosity integration or increase of the upper limit for the redshift integration, has no significant effect on the number counts $N(>S)$ in our flux range of interest. We note that all experimental XLF determinations are based on AGN samples, that cover a smaller luminosity range, typically $L \geq 10^{42} \text{ erg s}^{-1}$, and a smaller redshift range ($z_{\max} \approx 3 - 5$). Hence, our calculations involve some extrapolation of the measured XLFs to lower luminosities and higher redshift. The uncertainties introduced by this extrapolation are generally small, with a few exceptions that are discussed below.

Using the XLF of H05, we predict a somewhat smaller number of AGN for the soft-band counts than that observed by K07 and G08, with a strongest deviation of about 30 – 50 % for the $\log N - \log S$ curve in our flux range of interest. Part of this

¹² The solid angle is converted from steradian to square degrees ($\pi^2 \text{ sr} = 180^2 \text{ deg}^2$).

discrepancy arises because H05 selected only Type 1 AGN, and part may be caused by cosmic variance. It is beyond the scope of this work to investigate the origin of this difference in detail, therefore we renormalized the soft-band XLF of H05 upward with a factor of ≈ 1.35 to match the source counts of K07 and G08 in the flux range of interest. The hard-band $\log N - \log S$ obtained using the XLF of A10 agrees well with the observed source counts, with a strongest deviation of about 3 – 11 % in the flux range of interest.

2.5 Luminosity and redshift distribution of detected AGN

With the knowledge of the XLF (Sect. 2.4) we can compute luminosity and redshift distributions of detected AGN as follows:

$$\frac{dN(L)}{d \log L} = \int_0^{z_{\max}(S,L)} \phi(\log L, z) \frac{dV(z)}{dz} dz \quad (2.10)$$

$$\frac{dN(z)}{dz} = \frac{dV(z)}{dz} \int_{\log L_{\min}(S,z)}^{\log L_{\max}} \phi(\log L, z) d \log L \quad , \quad (2.11)$$

where z_{\max} is defined by the relation $d_L(z_{\max}) = \sqrt{L/(4\pi S)}$, where $d_L(z_{\max})$ is the luminosity distance at the redshift z_{\max} . For the other quantities and the K-correction see the explanation after Eq. (2.9). The corresponding cumulative distributions are

$$N(>L) = \int_L^{L_{\max}} dN(L') \quad (2.12)$$

$$N(>z) = \int_z^{z_{\max}} dN(z') \quad (2.13)$$

The number of AGN detected in the eRASS as a function of luminosity and redshift is summarized in Figs. 2.5 and 2.6 and is discussed in more detail in the next two subsections. In computing these distributions we took into account the sensitivity map (Sect. 2.2.4) of the eRASS via the sensitivity – solid angle distribution shown in Fig. 2.3. For the overexposed areas at the ecliptic poles we used the sensitivity quoted in Table 2.2. The properties of the brightest and faintest 10 % we computed using the flux limits from Sect. 2.3.

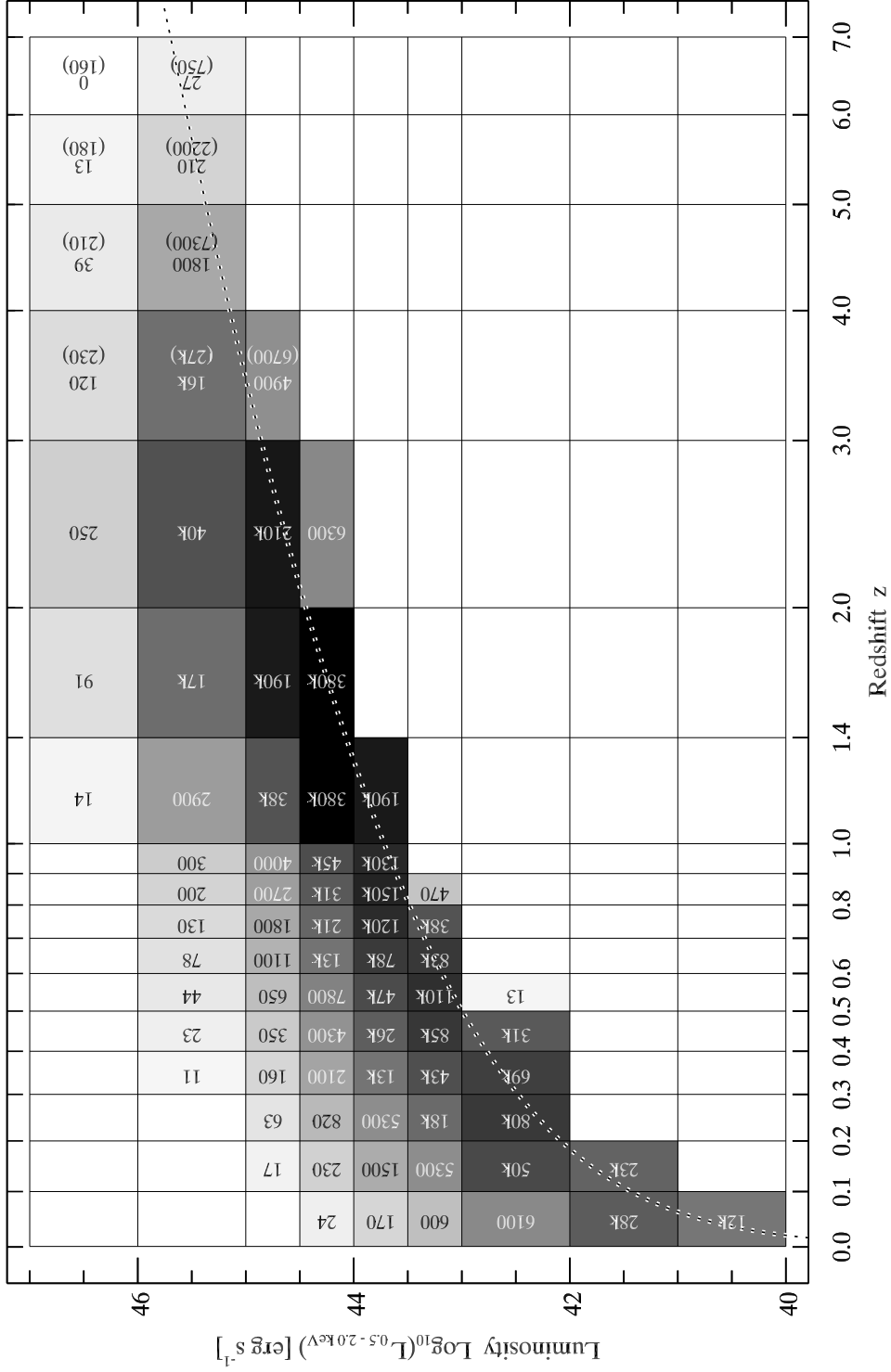


Figure 2.5: Number of AGN in different redshift and luminosity bins expected to be detected in the course of the four-year survey in the soft band ($0.5 - 2.0 \text{ keV}$). White empty bins with no number correspond to zero sources. The dotted line corresponds to the detection limit of eRASS. In the soft-band plot, the numbers in brackets are for the XLF without the exponential redshift cutoff, they are given only if the difference exceeds 10 %.

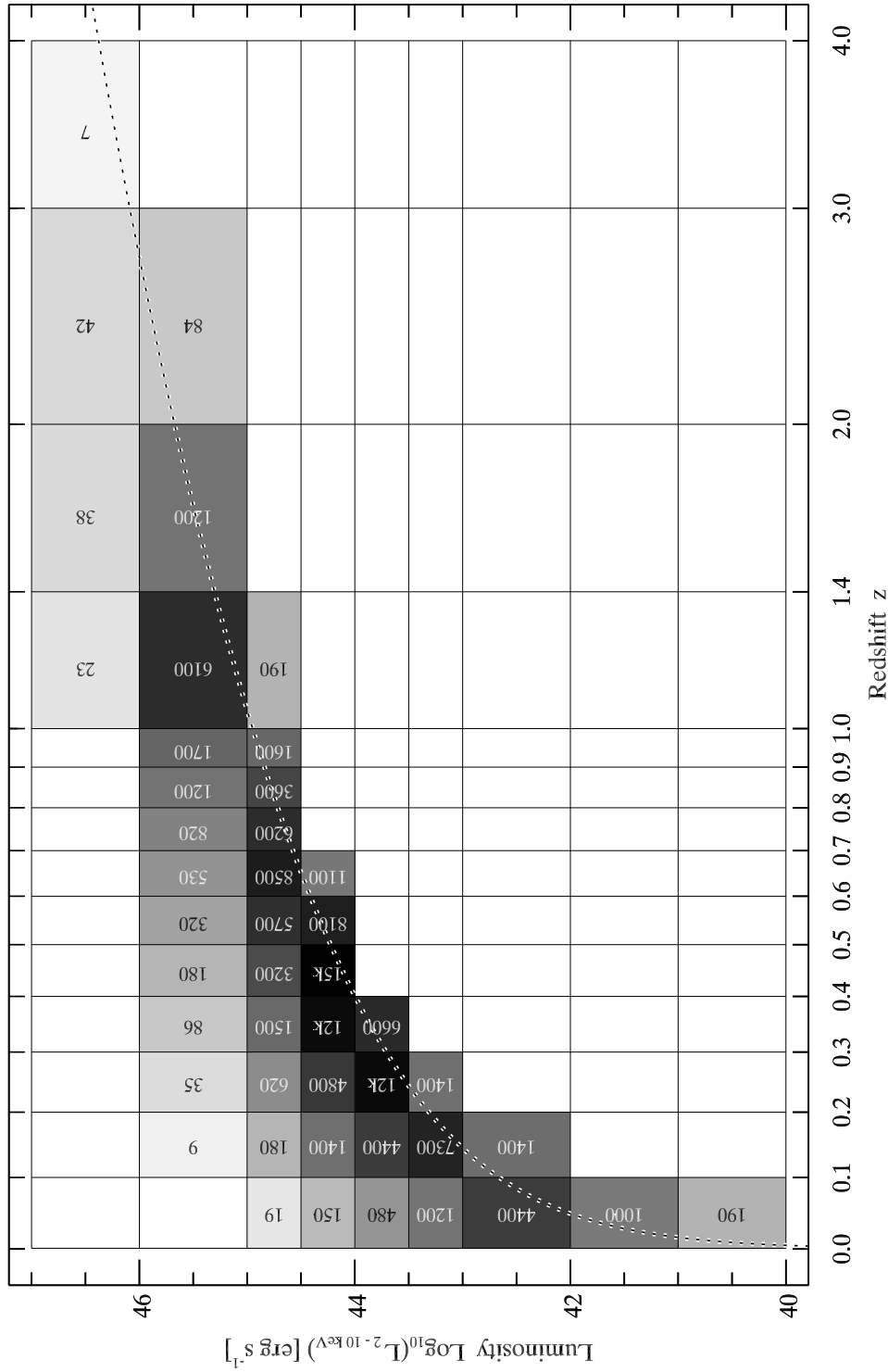


Figure 2.6: Same as Fig. 2.5, but for the hard band (2.0 – 10.0 keV).

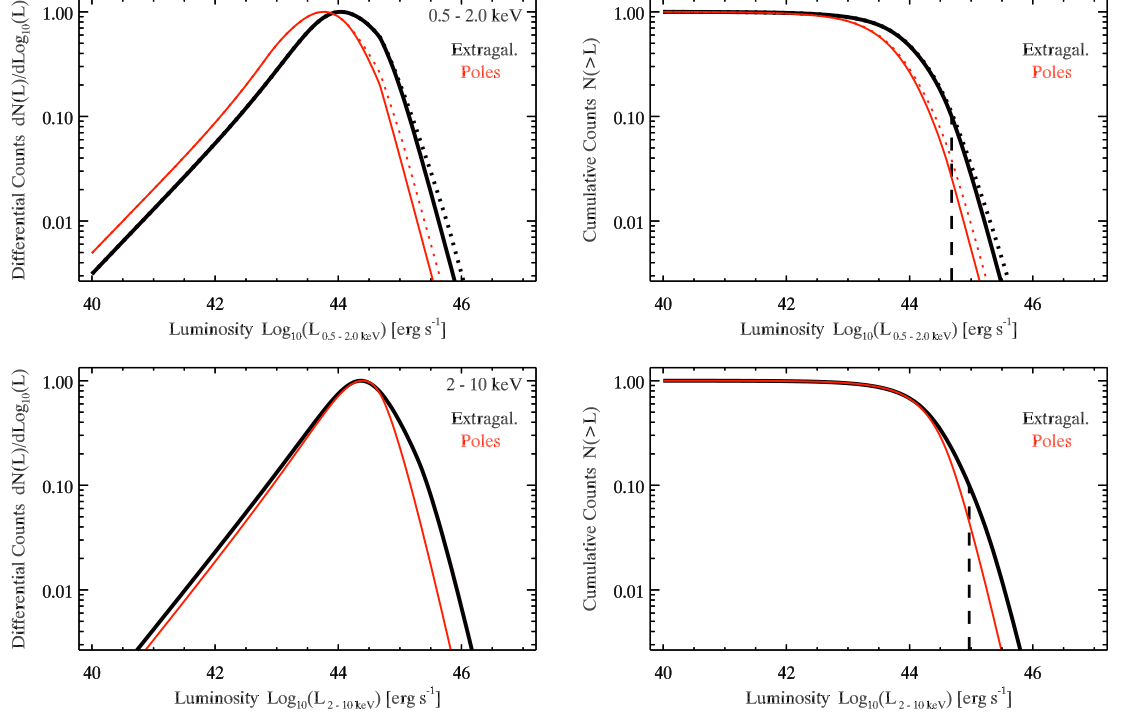


Figure 2.7: Differential (*left*) and cumulative (*right*) luminosity distribution in the soft (*top*) and the hard (*bottom*) band for the four-year eRASS AGN sample in the extragalactic sky (*black*) and at the ecliptic poles only (*red*). The distributions are normalized to unity to facilitate comparison of the shapes. The dotted curves in the top panels were computed without the high-redshift cutoff in the soft-band XLF (see Sect. 2.4). The dashed black vertical lines in the right panels show the luminosity corresponding to the 10 % fraction of sources.

2.5.1 Luminosity distribution

Luminosity distributions of detected AGN are shown in Fig. 2.7. In the soft band they peak at $\sim 10^{44}$ erg s $^{-1}$, with only a small difference between the extragalactic sky sample ($\approx 10^{44.0}$ erg s $^{-1}$) and the ecliptic poles ($\approx 10^{43.8}$ erg s $^{-1}$). The peak in the hard band occurs at $\approx 10^{44.4}$ erg s $^{-1}$. The median values differ by less than 1 % from the corresponding peak values. Comparing these with the values of L_* from the Table 2.3, the location of the peak is defined by the L_* luminosity and does not strongly depend on the survey sensitivity. A change of the latter by two orders of magnitude changes the position of the peak only by ~ 0.5 dex. Hence, our predictions for the luminosity distribution are very robust against moderate changes of the survey sensitivity. From the top panel of Fig. 2.7 one can see that the luminosity distribution in the soft band changes only marginally at high luminosity ($\gtrsim 10^{44}$ erg s $^{-1}$) if we exclude the high-redshift cutoff of the XLF.

From the cumulative luminosity distributions of the extragalactic sky (right panels of Fig. 2.7) about 10 % (vertical dashed lines) of the detected AGN will have luminosities higher than $\sim 10^{45}$ erg s $^{-1}$. This large sample of luminous AGN ($\sim 3 \times 10^5$ in full redshift range) will improve the constraints on the high-luminosity end of the XLF. For comparison, the AGN sample of H05 had about ~ 100 AGN with a luminosity higher than $\sim 10^{45}$ erg s $^{-1}$.

In the top panel of Fig. 2.9 the luminosity distribution of the brightest 10 % (those with the highest flux, blue curve) and the faintest 10 % (lowest flux, red) do not differ much from each other and from the distribution of the total sample (black). We note that the luminosity distribution of objects detected in the half-year survey is well represented by that of the brightest 10 % of the sources.

2.5.2 Redshift distribution

Unlike the luminosity distributions, redshift distributions of a flux-limited sample are strongly dependent on the limiting flux (Fig. 2.8). Correspondingly, the redshift distributions for the extragalactic sky sample and for the poles peak at different redshift, the difference being larger for the hard band. The same is true for the median values, which are listed together with the peak values in Table 2.4. The median and peak values in the soft band do not change significantly when we exclude the exponential high-redshift cutoff from our calculations. However, the redshift distribution of AGN at high-redshift does change significantly, which we can see clearly in the top panel of Fig. 2.8. This is discussed in more detail in Sect. 2.5.4.

The differential distributions show several breaks that are caused by the derivative discontinuities of the LDDE model. Another break at the low redshift, $z \sim 0$, appears

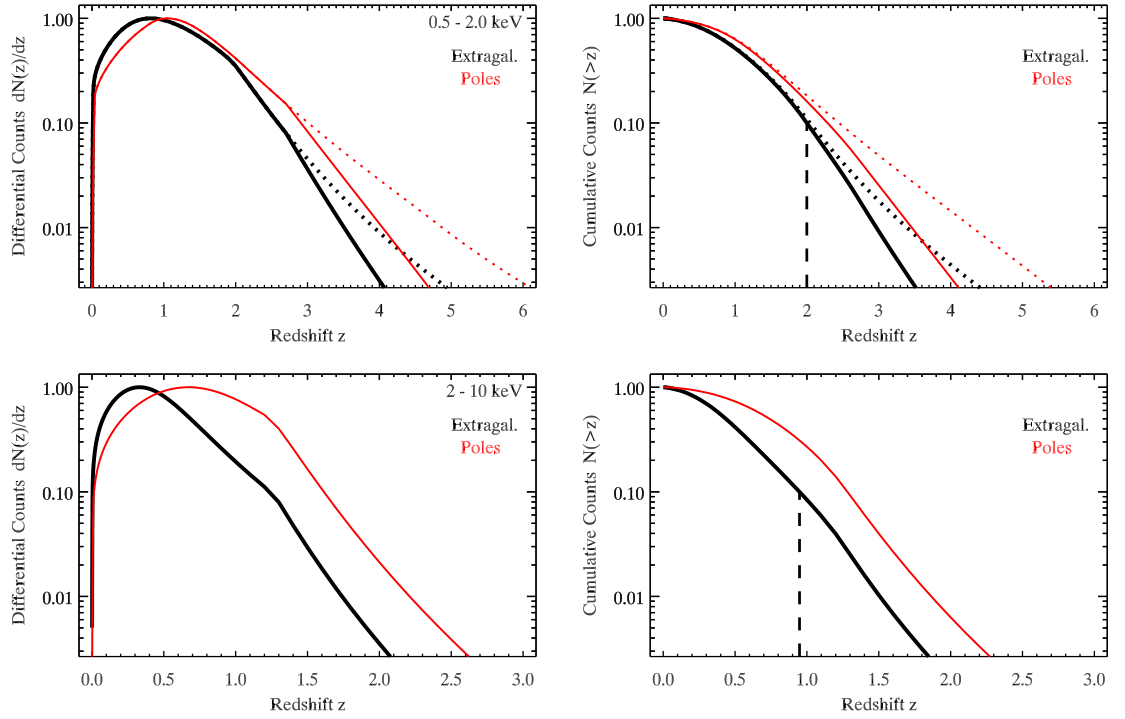


Figure 2.8: Same as Fig. 2.7, but for the redshift distribution.

Table 2.4: Peak and median values of the redshift distribution of eRASS AGN

	0.5 – 2.0 keV		2.0 – 10.0 keV	
	Peak	Median	Peak	Median
4.0 years extragalactic	0.8	1.0	0.3	0.4
4.0 years ecliptic poles	1.0	1.2	0.7	0.7
Brightest 10 %	0.3	0.5	0.2	0.2
Faintest 10 %	1.1	1.2	0.5	0.5
0.5 years extragalactic	0.4	0.6	0.2	0.2

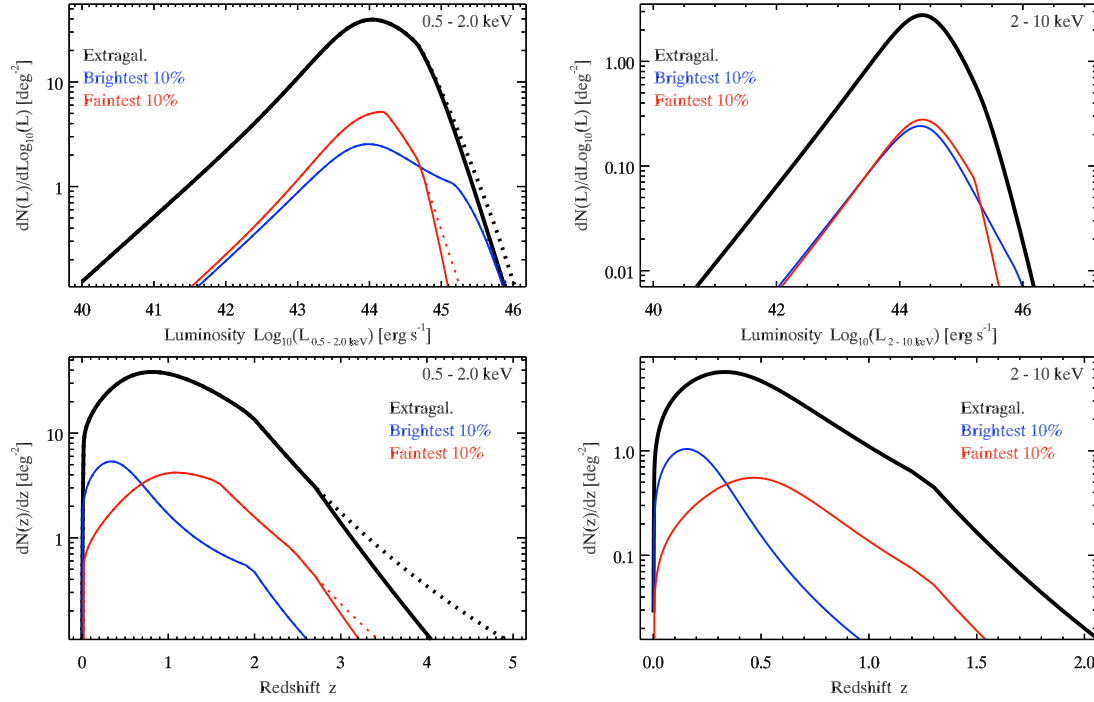


Figure 2.9: Differential luminosity (*top*) and redshift (*bottom*) distributions in the soft (*left*) and hard (*right*) bands for the entire extragalactic AGN sample after four years (*black*) and for the brightest 10% (*blue*) and faintest 10% (*red*). The black curves are same as in Figs. 2.7 and 2.8, but without renormalization to unity.

when the low integration limit in Eq. (2.11), $\log L_{\min}(S, z)$, becomes equal to the low limit of the interval where the XLF is defined ($L = 10^{40.0} \text{ erg s}^{-1}$, Sect. 2.4). These features are not physical and reflect the deficiencies of the functional form used in the LDDE model. However, these deficiencies of the XLF model do not compromise the overall shapes of the redshift (and luminosity) distributions derived in this paper, as long as the overall shape of the AGN X-ray luminosity function is adequately represented by the LDDE model.

In accord with the note made at the beginning of this Section, redshift distributions of the brightest and faintest 10 % of the AGN (Fig. 2.9) peak at significantly different redshifts than the overall distributions (black). Similar to luminosity distributions, the redshift distributions of the objects detected during the half-year survey are similar to the distributions of the brightest 10 %. From the cumulative distributions (right panels in Fig. 2.8.) we conclude that in the soft band, $\approx 50\%$ of objects in the eRASS sample will be located at $z > 1$, whereas $\approx 10\%$ will be located at $z > 2$.

To illustrate the potential of the eRASS AGN sample in the limited sky areas, we show in Fig. 2.10 the number of objects per redshift bin as a function of redshift for several luminosity groups. For this calculation we chose a sky area of $14\,000 \text{ deg}^2$, similar to the area of the Sloan Digital Sky Survey (SDSS), and considered relatively broad redshift bins, consistent with the expected accuracy of photometric redshifts based on the multiband photometry (Salvato et al. 2011). It is obvious from Fig. 2.10 that even coarse redshift information over relatively limited areas of sky is capable of delivering unprecedented samples of AGN, covering the most luminous AGN ($> 10^{45} \text{ erg s}^{-1}$) with an unmatched statistical significance.

2.5.3 Uncertainties

Obviously, the accuracy of our predictions depends on the accuracy of the AGN XLF. This is limited by the moderate numbers of objects used for their construction, typically about 1 000. Although the XLFs obtained by different authors are broadly consistent with each other, there is still a considerable spread between different models. Correspondingly, using XLF models obtained by different authors, we obtained somewhat varying luminosity and redshift distributions of eRASS AGN.

To illustrate the range of uncertainties, we calculated the luminosity and redshift distributions for the soft-band extragalactic sky sample using several different XLF models available in the literature. Along with our default soft-band XLF, we used the XLF of H05 without the exponential redshift cutoff and the XLF models of Miyaji et al. (2000, Table 3) and Ebrero et al. (2009). These XLF models are based

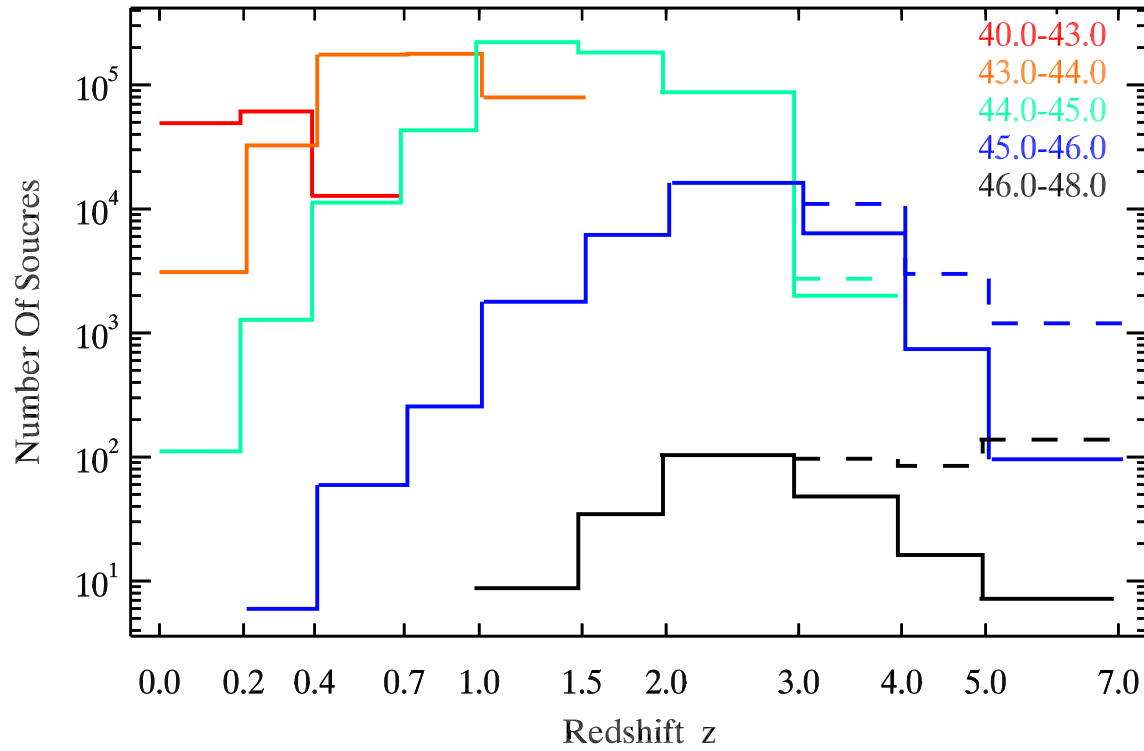


Figure 2.10: Number of eRASS AGN in the soft band (0.5–2.0 keV) as a function of the redshift for different luminosity groups in a sky area similar to that covered by SDSS (14 000 deg²). The solid and dashed histograms show predictions based on our default soft-band XLF with and without exponential high-redshift cutoff (Sect. 2.4), respectively.

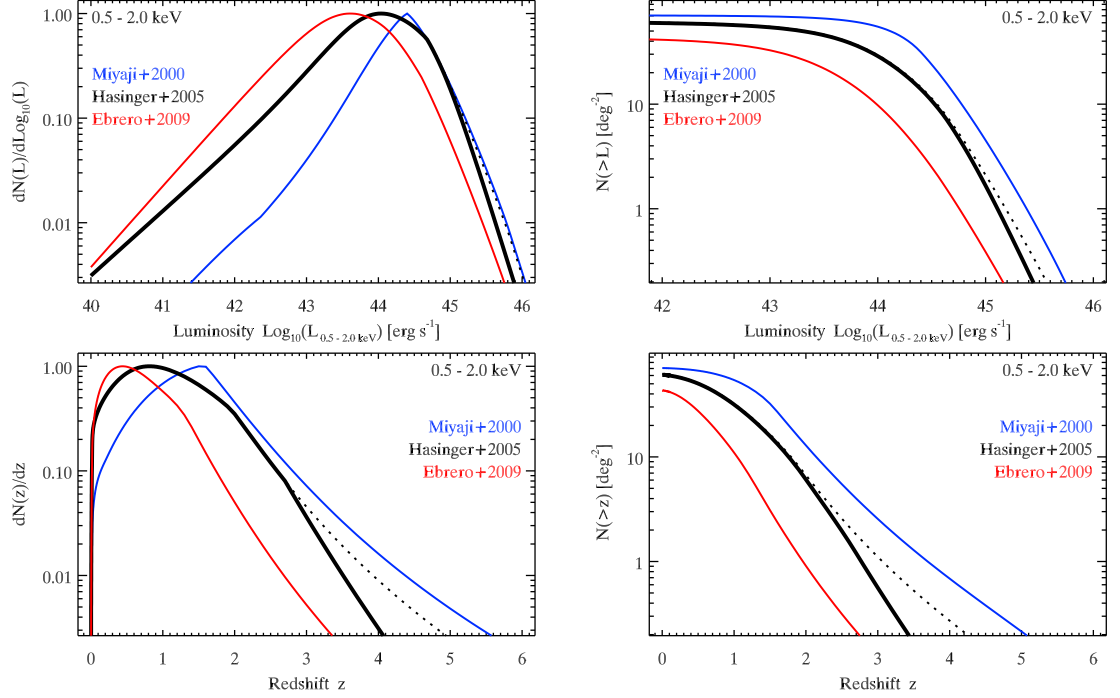


Figure 2.11: Differential (*left*) and cumulative (*right*) luminosity (*top*) and redshift (*bottom*) distributions of the soft-band extragalactic sky sample computed using different XLFs. The thick solid black curves show predictions based on the default XLF model. Also shown are predictions for the XLF model of Miyaji et al. (2000, Table 3), Hasinger et al. (2005, Table 5, without the exponential redshift cutoff, dotted curve), and Ebrero et al. (2009). To facilitate the comparison of shapes, differential distributions in left panels are normalized to unity.

on (overlapping) samples, each containing about 1 000 objects in total. Because the samples partly overlap, the models are not entirely independent. The resulting luminosity and redshift distributions of the eRASS extragalactic sky sample are shown in Fig. 2.11. To facilitate the comparison of the shapes, the differential distributions are normalized to unity, whereas the cumulative distributions are shown with their original normalization.

It can be seen from Fig. 2.11 that different XLFs predict considerably different luminosity and redshift distributions for eRASS AGN. Although the detailed shapes may be not critically important for the purpose of this paper, median redshifts and luminosities are important characteristics of the eRASS AGN sample. These parameters differ somewhat significantly for the three XLFs, with Miyaji et al. (2000) and Ebrero et al. (2009) presenting two extremes and our default XLF of Hasinger et al. (2005) located in the middle. It is worth mentioning that the primary reason for selecting the XLF of Hasinger et al. (2005) as our default one was the fact that it best reproduced the observed $\log N - \log S$ distributions (Sect. 2.4). Note also that the total numbers of eRASS AGN are predicted sufficiently accurately from the observed $\log N - \log S$ distributions and therefore are not of significant concern.

In interpreting the plots in Fig. 2.11 one should keep in mind, that different authors applied different selection criteria in building their samples and used slightly different versions of the LDDE model. In addition, Ebrero et al. (2009) applied a correction for absorption, which the others did not. These differences explain, in particular, the difference in the total AGN surface density predicted by different models. They also explain, at least in part, the considerably large discrepancy in the shapes of the predicted luminosity and redshift distributions. Another part of the difference is probably related to statistical uncertainties in the XLF parameters. Although a detailed comparison of XLFs produced in different studies is beyond the scope of this paper, it would be interesting to see to which extent the discrepancy can be explained by statistical uncertainties. However, the LDDE model is a multi-parameter model with a complex correlation pattern between parameters. Therefore a proper error analysis would require knowledge of the covariance matrixes, which are not available anymore (T. Miyaji, priv. comm.). An attempt to treat the XLF parameter errors as independent resulted in unreasonably large uncertainties in the predicted distributions for the eRASS sources. On the other hand, the moderate size of the statistical errors in the observed XLF data in the redshift and luminosity range relevant to the bulk of eRASS AGN ($L \sim 10^{44} \text{ erg s}^{-1}$, $z \sim 1$) suggests that the behavior of the derived distributions near their peaks is probably not significantly affected by the propagation of statistical errors, therefore the differences seen in Fig. 2.11 probably reflect genuine differences in XLFs.

Another important factor to be taken into consideration is cosmic variance. Because the AGN XLF determinations rely on the survey that covers only a small fraction of the sky, $\lesssim 10^{-4} - 10^{-3}$ at most, they are subject to cosmic variance. The amplitude of this uncertainty is probably in the $\sim 10\%$ range (Aird et al. 2010). Obviously, the eRASS sample will provide means for studying this effect in full detail.

We emphasize that we did not consider any separation between different types of AGN. H05 only considered type 1 AGN for their XLF model. If we take into account the expected small fraction ($\sim 10\%$ ¹³, see also the SB) of type 2 AGN (intrinsic $N_{\text{H}} > 10^{21} \text{ cm}^{-2}$) and the fairly similar XLF of both types (e.g. Burlon et al. 2011), we expect that the introduced uncertainties will be relatively small.

Finally, we did not take into account the Eddington bias, neither did we consider the details of the source detection and background subtraction algorithms, which will affect to some extent the numbers of detected sources and their $\log N - \log S$ distributions at the faint end. They will also affect the completeness characteristics of the eRASS AGN sample, which will have to be accounted for in constructing XLFs. These are typical properties of flux-limited surveys, especially those conducted in the photon-counting regime, in the limit of small numbers of counts, where the character of the Poissonian distribution of counts manifests itself strongly. The data analysis methods and techniques used to deal with these effects are well known and constitute the standard set of tools in X-ray astronomy. A detailed account of these effects and others (e.g. confusion with extended sources) is beyond the scope of this paper.

2.5.4 High-redshift AGN

The density of AGN at high-redshifts is of particular interest because it can place constraints on the formation scenarios of first supermassive black holes and, hence, on cosmological models (Brandt & Hasinger 2005). Their numbers in the existing surveys, including those used to produce the XLF models, are very limited. Indeed, the highest redshift bin in the AGN sample of H05 was located at $z = 3.2 - 4.8$ and contained 17 objects. The sample of Miyaji et al. (2000) contained 25 AGN in a somewhat wider redshift interval of $2.3 - 4.6$, and the sample of Ebrero et al. (2009) had no AGN with $z > 3$. Moreover, these samples are not entirely independent because they were obtained from overlapping sets of deep surveys. There is only a handful of $z > 5$ AGN currently known in X-rays (e.g. Civano et al. 2011). Due to low numbers of distant AGN, the XLF at high redshifts is poorly constrained. As demonstrated below, eRASS will significantly enhance the statistics of high-redshift objects.

¹³<http://www.bo.astro.it/~gilli/counts.html>

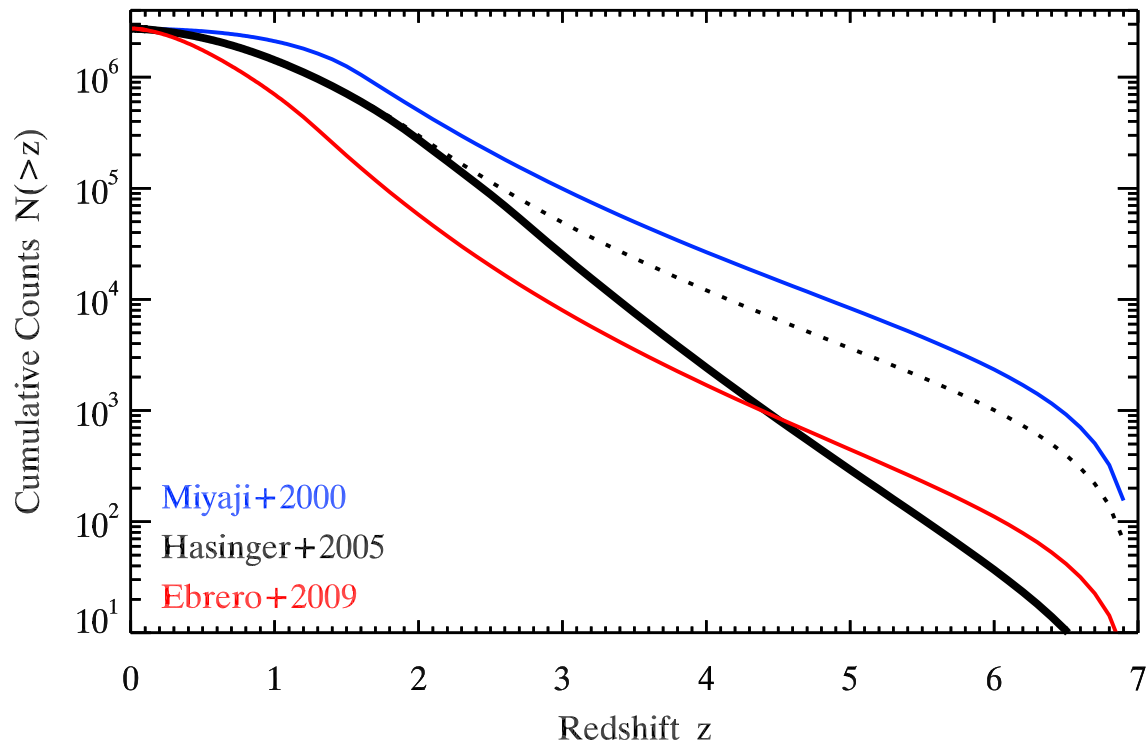


Figure 2.12: Numbers of high-redshift AGN, $N(> z)$, expected in the soft band ($0.5 - 2.0$ keV) for the extragalactic sky after four years. The thick solid black curve shows the prediction based on the default XLF model. Also shown are predictions for the XLF model of Miyaji et al. (2000, Table 3), Hasinger et al. (2005, Table 5, without the exponential redshift cutoff, dotted curve), and Ebrero et al. (2009). To obtain these curves we integrated the XLFs to the highest redshift of $z = 7$. All curves are rescaled to match the average source density computed with our default model (Sect. 2.3).

Our poor knowledge of the AGN XLF at high-redshifts limits our ability to accurately predict numbers of high-redshift AGN in the eRASS. To estimate the range of uncertainties we calculated their numbers in the soft band using several different XLF models. The resulting cumulative number counts are shown in Fig. 2.12. For the purpose of this comparison, the curves were rescaled to reproduce the same number density of AGN as the arithmetic mean of our number density map introduced in Sect. 2.3. The correction factors in the soft band are 1.33 and 1.32 for H05 with and without redshift cutoff, respectively, 1.15 for Miyaji et al. (2000, Table 3) and 1.88 for Ebrero et al. (2009).

As one can see from the Fig. 2.12, the number of high-redshift objects in the extragalactic sky sample is uncertain by more than an order of magnitude, different predictions ranging for $z > 3$ from $\sim 10^4$ to $\sim 10^5$. For $z > 4$ and $z > 5$ the numbers vary from ~ 2000 to ~ 30000 and from ~ 300 to ~ 9000 . The exponential redshift cutoff of the H05 XLF (cf. solid and dotted black curves) has a significant effect on the numbers of high-redshift sources bringing it close to the prediction based on the XLF of Ebrero et al. (2009, red curve). On the other hand, the prediction based on the XLF of H05 without a cutoff (dotted black curve) is close to that of Miyaji et al. (2000, blue curve). This large discrepancy of different XLF at high-redshift was previously pointed out by Brusa et al. (2011).

For the number of objects in the $z = 3 - 5$ redshift range our predictions vary from ~ 8000 to ~ 90000 , the default XLF giving ~ 30000 objects; without the exponential redshift cutoff this number is increased by almost a factor of two.

According to our default hard-band XLF, there will be ~ 4 detected AGN in the hard band for the redshift of $z \gtrsim 3.5$. However, the discrepancy between different hard-band XLFs is also large, the predictions ranging from ~ 2 (Aird et al. 2010, LADE model of Table 4) to ~ 200 (La Franca et al. 2005) AGN for $z \gtrsim 3.5$.

The density of high-redshift objects will be higher in the ecliptic poles (Fig. 2.8 and Table 2.2). For the default soft-band XLF, there will be 1 high-redshift ($z \gtrsim 5$) AGN every $\sim 5 \text{ deg}^2$, which results in ~ 17 objects in total. Without the exponential redshift cutoff this number is an order of magnitude larger. For the other rescaled XLFs from Fig. 2.12 the number of objects varies between ~ 20 and ~ 600 . Obviously, the higher source density and smaller area will facilitate the search for high-redshift objects in the pole regions.

2.6 Redshift determination with the iron $K\alpha$ line

The strong $K\alpha$ line of iron at $\approx 6.4 \text{ keV}$ in the spectra of AGN in principle opens the possibility of determining redshifts by means of X-ray spectroscopy. Below,

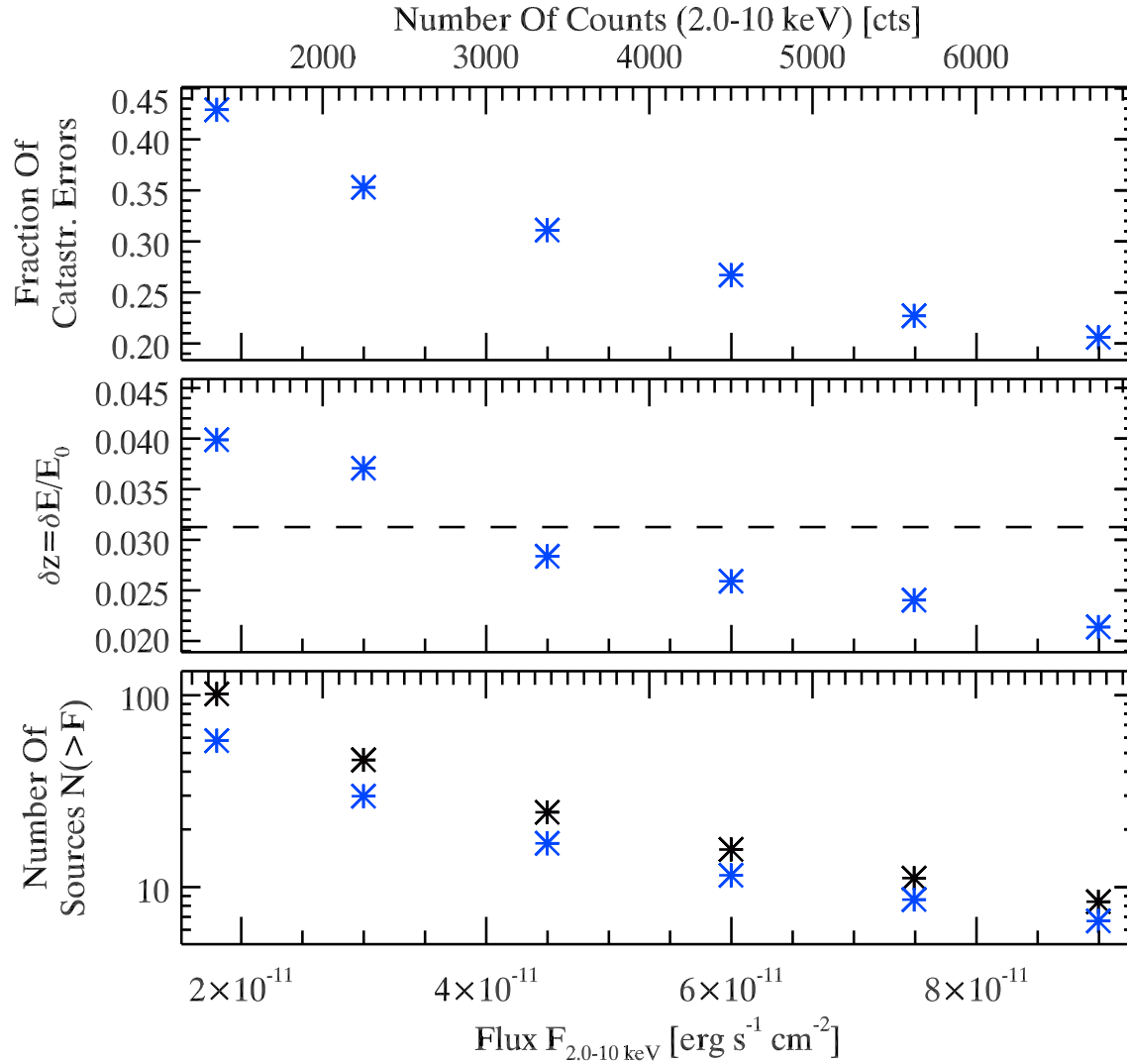


Figure 2.13: Feasibility of using the iron $K\alpha$ line for the redshift determination at $z = 0$. See Sect. 2.6 for details. *Top*: Fraction of catastrophic failures. *Middle*: Accuracy of the redshift determination. The black dashed line corresponds to the energy resolution of eROSITA. *Bottom*: Number of sources for which the redshift can be determined with the accuracy shown in the middle panel, the catastrophic failures excluded (black points show only the $\log N - \log S$).

we investigate this possibility for the parameters of eRASS characteristics of the eROSITA telescope.

We assumed that the continuum spectrum is described by an absorbed power-law with a photon index of $\Gamma = 1.9$ and $N_{\text{H}} = 6 \times 10^{20} \text{ cm}^{-2}$. The shape of the iron $\text{K}\alpha$ line may be fairly complex and typically includes narrow and broad components, with centroids depending on the ionization state (e.g. 6.4 keV for the neutral component), and may be complicated by other features (e.g. the 7.11 keV absorption edge; e.g. Gilli et al. 1999; Nandra et al. 2007; Corral et al. 2008; Shu et al. 2010; Krumpe et al. 2010a; Chaudhary et al. 2012). However, as the final result of this calculation turn out to be somewhat negative, we ignored this complexity and used a simple model¹⁴, consisting of a single Gaussian line at 6.4 keV with an intrinsic width of $\sigma_{\text{Fe}} = 200 \text{ eV}$ and an equivalent width of 150 eV (rest-frame values).

To investigate the detectability of the iron $\text{K}\alpha$ line in the spectra of eRASS AGN, we performed the following simulations: We chose a number of flux values in the $10^{-11} - 10^{-10} \text{ erg s}^{-1} \text{ cm}^{-2}$ flux range. We fixed the redshift and for each flux value simulated 1000 spectra using the `phabs(zpowerlw + zgauss)` model in XSPEC (version 12.7.0, Arnaud 1996). Each spectrum was fit with the same model. In the fit, the parameters σ_{Fe} and N_{H} were fixed, the initial values of other parameters were set at their simulated values. After 1000 spectra were simulated, the distribution of the best values of the redshift was analyzed. It was fit with a Gaussian distribution, then the points outside the $\pm 3\sigma$ range were marked as catastrophic failures and clipped out, and the distribution was fit by a Gaussian again. The newly obtained width of the Gaussian determines the accuracy of the redshift determination σ_z . The catastrophic error fraction was then recomputed as a fraction of objects outside $\pm 3\sigma_z$.

Our results for the redshift $z = 0$ are shown in Fig. 2.13. As one can see from the plot, even at considerably large number of counts, ~ 1500 in the hard band, corresponding to a flux of $\sim 2 \times 10^{-11} \text{ erg s}^{-1} \text{ cm}^{-2}$, the fraction of catastrophic errors is still large, $\sim 40\%$. This is caused by the steep decrease of the eROSITA efficiency curve with energy, by more than an order of magnitude between 2 and 6 keV. As a result, even at large total number of counts, the number of counts at $\sim 6 \text{ keV}$, is too small for a reliable line detection in the flux range of interest.

From the middle panel of the Fig.2.13, the accuracy of the redshift measurements for the remaining $\sim 60\%$ of objects may seem to be reasonably good, $\delta z \lesssim 0.05$. Obviously, this is a result of its definition, which relies on excluding catastrophic failures. This definition works well when there are few catastrophic failures. However, the effect of small δz is nullified when the fraction of catastrophic failures

¹⁴ In the flux range of interest, more complex models generally lead to larger fraction of catastrophic failures.

is large. Furthermore, the numbers of objects in this flux range is of the order of 100 on the entire extragalactic sky, which is too small to be useful. The majority, if not all of these bright objects, will be previously known AGN with known redshifts.

The increase of the effective area toward low energies could improve the situation at higher redshifts. For a $z \approx 2$ object, for example, the observed energy of the iron $K\alpha$ line would fall near the peak of the eROSITA sensitivity and would lower the lowest flux required for a reliable redshift determination using the iron $K\alpha$ line to $\sim 10^{-12} \text{ erg s}^{-1} \text{ cm}^{-2}$. Unfortunately, the relatively small number of $z \approx 2$ objects (a few hundreds) and their low fluxes negate the advantage given by the larger effective area at low energies. However, we note that a spectral analysis would still be possible for the sources of known redshift, and that one can still use spectral stacking analysis (e.g. Chaudhary et al. 2010) to study the average properties of the iron $K\alpha$ line of AGN.

2.7 Optical counterparts

To fully explore the potential of the eRASS, extensive optical coverage will be critical. The primary, but not sole goal of this coverage will be to provide identifications and redshift information. A detailed discussion of the feasibility and possible strategies of the optical support of the eRASS and its synergies with various ongoing and future optical surveys is beyond the scope of this paper and is currently extensively discussed in the eROSITA collaboration. In this Section we investigate the expected optical magnitude distribution of the eRASS AGN. To this end, we used results of the XMM-COSMOS (Hasinger et al. 2007) and XBootes (Murray et al. 2005) surveys. For the COSMOS field, we used the results of Brusa et al. (2010), who cross-correlated the original XMM-COSMOS catalog of X-ray sources of Cappelluti et al. (2009) with the data of the optical survey of the COSMOS field by Capak et al. (2007). From these data we selected sources with a $0.5 - 2.0 \text{ keV}$ flux, which exceeds the eRASS four-year detection threshold (214 out of 1710 sources fulfill the condition $S_{0.5-2.0 \text{ keV}} \geq 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$), which had a high identification reliability (204 out of 214), and were not brighter than $I_{\text{AB}} = 14.0 \text{ mag}$ (197 out of 204). This selection resulted in a sample of 197 sources for which we obtained the I -band AB-magnitude distribution.

We similarly analyzed the XBootes field, cross-correlating the X-ray and optical catalogs for this field (Kenter et al. 2005; Brand et al. 2006). We selected point-like sources (class ≥ 0.50) with $S_{0.5-2.0 \text{ keV}} \geq 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$ (565 out of 3213), for which an optical counterpart was found (St = 1) (565 out of 565) with a high probability of true identification (Popt ≥ 0.95) (561 out of 565) and an optical flux

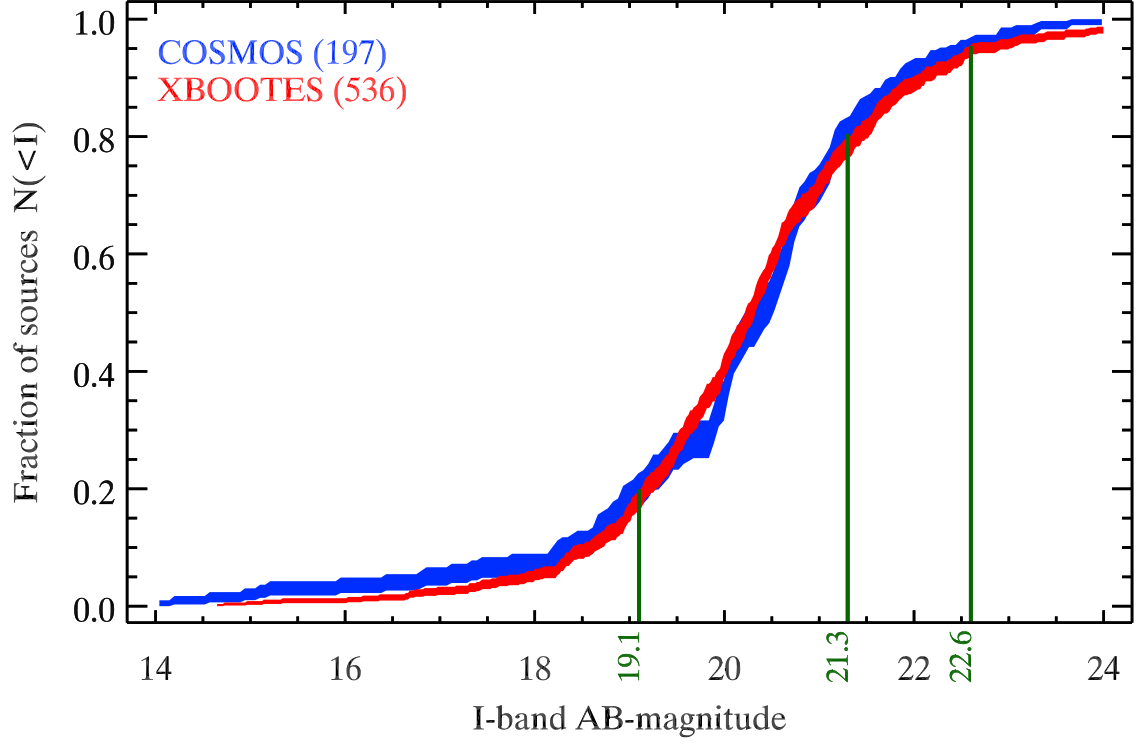


Figure 2.14: Cumulative I -band AB-magnitude distribution of AGN in the COSMOS and XBOOTES fields with the $0.5 - 2.0$ keV flux exceeding the four-year eRASS detection threshold. The thickness of the curves represents the standard deviation of a binomial distribution. The vertical lines show the photometric sensitivities of the SDSS (21.3 mag) and Pan-STARRS PS1 (22.6 mag) and the magnitude limit for SDSS spectroscopy (19.1 mag).

of $I_{AB} > 14.0$ mag (536 out of 561). We thus selected 536 X-ray sources, for which we computed the cumulative I -band magnitude distribution, converting the Vega magnitudes to AB-magnitudes with the conversion factor from Blanton & Roweis (2007, Table 1): $m_{AB} = m_{Vega} + m_{AB}(Vega)$ with $m_{AB}(Vega) = 0.45$ for the I -band.

These I -band magnitude distributions for the COSMOS and Bootes fields are plotted in the Fig.2.14. They show a good agreement between the results for the two different fields, meaning that we have a very good knowledge of the expected magnitude distribution of sources at bright X-ray fluxes. Comparing this distribution with the limiting magnitude of the Sloan Digital Sky Survey in the i -band, 21.3 mag at the 95 % completeness (Abazajian et al. 2009), we conclude that about $\approx 80\%$ of the eRASS AGN in the SDSS sky will have optical counterparts. Taking into account the sky area covered by SDSS, $\approx 14\,500$ deg², we estimate that about $\sim 1/3$ of eRASS objects will have an optical counterpart in the SDSS photometric catalog. About $\approx 20\%$ of objects will be brighter than the spectroscopic limit of the SDSS, $i = 19.1$ mag for quasars at $z < 3$ (Richards et al. 2002), that is, some fraction of these objects will have SDSS spectra.

Repeating this analysis for the half-year sensitivity of eRASS, we expect that almost all eRASS AGN in the SDSS sky will have optical counterparts.

One can see from Fig. 2.14 that approximately 95 % of eRASS AGN will be brighter than $I_{AB} = 22.5$ mag ($R_{AB} \approx 23.0$ mag). The Pan-STARRS PS1 3π survey will exceed this depth with its expected sensitivity of ≈ 22.6 mag in one visit (Chambers & the Pan-STARRS team 2006). The three year PS1 sensitivity in the I -band will reach ≈ 23.9 mag and will cover virtually all eRASS objects in the field of the 3π survey.

2.8 Summary and conclusions

We computed various statistical characteristics of the expected eRASS AGN sample, including their luminosity- and redshift distributions, and the magnitude distributions of their optical counterparts.

The eROSITA all-sky survey will produce an unprecedented sample of about 3 million X-ray selected AGN. With a median redshift of $z \approx 1$, approximately 40 % of the eRASS objects will be located between redshifts $z = 1$ and $z = 2$ (Fig. 2.8). We predict that about $10^4 - 10^5$ AGN beyond redshift $z = 3$ and about 2 000 – 30 000 AGN beyond redshift $z = 4$, the exact numbers depending on the behavior of the AGN XLF in the high-redshift and high-luminosity regimes (Fig. 2.11).

The eRASS AGN sample will open the possibility of studying the growth of super-massive black holes, their co-evolution with host galaxies and dark matter halos, and

their relation to the large-scale structure to unprecedented detail, and potentially, it may also help to constrain cosmological parameters (Chapter 3 Htsi et al. 2014). Importantly, it will permit one to conduct these studies beyond redshift $z = 1$, which is poorly covered by the current optical surveys.

To fully exploit this potential of eRASS, an extensive optical support will be critical. One of the main goals of optical follow-up will be to provide redshifts for eRASS AGN, but its importance will reach far beyond this, including, for example, studies of the co-evolution of supermassive black holes and their host galaxies (see Sect. 3.5.3 and Htsi et al. (2014) for a discussion of the different goals of optical follow-ups and their requirements and prospects). With the capabilities of the currently available facilities and their time allocation strategies, measuring of optical spectra for the entire sample of the ~ 3 million objects does not appear to be achievable on realistic timescales. However upcoming hardware and survey programs and proposals, for instance, 4MOST (de Jong et al. 2012) and WEAVE (Dalton et al. 2012), can make this task more realistic, especially for some limited areas of sky. Furthermore, introducing multiband photometry and other improvements of the photometric redshift measurement techniques will make determining of photometric redshifts for large samples of eRASS AGN possible (Salvato et al. 2011; Saglia et al. 2012).

Chapter 3

Large-scale structure studies with AGN of the eROSITA all-sky survey

3.1 Introduction

Large-scale structure (LSS) studies are established as an important tool for studies in two major areas of astrophysics: cosmology, and galaxy evolution. A key of their success is the increasing number of surveys at different wavelengths with increasing depths and sky coverages. In X-rays, many deep, extragalactic surveys have been performed in the past decade (Brandt & Hasinger 2005; Cappelluti et al. 2012a; Krumpke et al. 2014). However, in comparison with other wavelengths, X-ray surveys with a large sky coverage and sufficient depth are still rare. The previous X-ray all-sky survey was performed by ROSAT¹ (Truemper 1993; Voges et al. 1999) more than two decades ago. Its successor with an ~ 30 times better sensitivity will be the four-year long all-sky survey (eRASS) of the eROSITA² telescope (Predehl et al. 2010), to be launched aboard the Russian Spektrum-Roentgen-Gamma satellite³ in 2016.

The major science goals of the eROSITA mission are studying cosmology with clusters of galaxies and active galactic nuclei (AGN) and constraining the nature of dark matter (DM) and dark energy. For a comprehensive description of the eROSITA mission we refer to the science book of eROSITA (Merloni et al. 2012).

¹<http://www2011.mpe.mpg.de/xray/wave/rosat/>

²<http://www.mpe.mpg.de/eROSITA>

³<http://hea.iki.rssi.ru/SRG>

In this work, we explore the potential of studying LSS with the AGN sample to be detected in eRASS. We focus on two important aspects of LSS studies: the clustering strength (represented by the linear bias factor, Sect. 3.3) and the baryon acoustic oscillations (BAOs, Sect. 3.4). To measure the former quantity, the redshift accuracy of photometric surveys is sufficient, therefore bias studies can be successfully conducted during and soon after the time eRASS is concluded. The BAO measurements, on the other hand, will be much more difficult to accomplish because spectroscopic redshift accuracy over large sky areas will be required. Note that a sufficient redshift accuracy can also be provided by high-quality narrow-band multifilter photometric surveys.

In our previous work (Chapter 2), we have studied the statistical properties of the AGN sample of eRASS and will adopt these results here. In the current work, we focus on the AGN detected in the soft-energy band ($0.5 - 2.0$ keV) and on the extragalactic sky ($|b| > 10^\circ$, $\sim 34\,100$ deg²). In the following calculations we assumed the four-year average sensitivity of 1.1×10^{-14} erg s⁻¹ cm⁻² adopted from Table 2.2 and luminosities are given in the soft band ($0.5 - 2.0$ keV).

Large optical follow-up surveys will be needed to provide identification and redshift information to the desired accuracy for all eRASS AGN. Current optical surveys are not sufficient in size and/or depth. A sensitivity of $I \approx 22.5$ mag ($R \approx 23.0$ mag) is required to detect at least 95% of the eRASS AGN (see Sect. 2.7 for details). Many photometric and spectroscopic surveys with different parameters have been proposed or are being currently constructed (e.g. Merloni et al. 2012). For the purpose of our investigation we assumed that redshifts are available for all eRASS AGN. We explore the effects of redshift-errors in our subsequent study (Hütsi et al. 2014).

3.2 Angular power spectrum

The commonly used tool for studying LSS is to measure and analyze the clustering of objects (such as AGN) with the 2-point correlation function (2pCF) or the power spectrum (e.g., Peebles 1980). These two methods, 2pCF and power spectrum, have their benefits and disadvantages (e.g. Wall & Jenkins 2012) but contain the same information about the LSS because they are related via Fourier transform. We used the angular power spectrum C_ℓ to characterize the clustering properties of objects. To predict the power spectra that will be measured with eRASS AGN, we relied on the model for AGN clustering of Hütsi et al. (2012). In particular, we used their model II, where they employed an observationally determined AGN X-ray luminosity function (XLF) and assumed that the linear bias factor of AGN corresponds to the

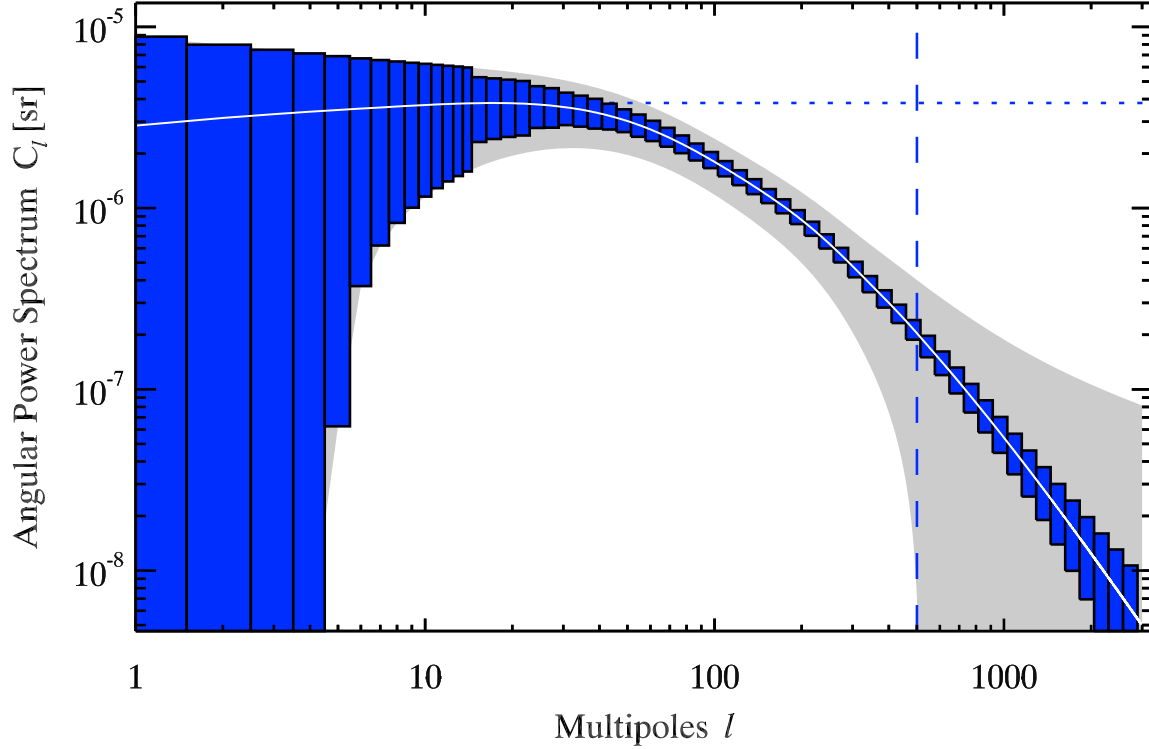


Figure 3.1: Angular power spectrum of the full eRASS AGN sample (after four years) in the soft band ($0.5 - 2.0$ keV) for the extragalactic sky ($f_{\text{sky}} \approx 0.83$) and $0 < z < 5$. The grayshaded area and the blue histogram show the 1σ uncertainty region (Eq. 3.4) without and with l -binning, respectively. The horizontal dotted line shows the level of shot noise, which was already subtracted from the angular power spectrum. For multipoles above the vertical dashed line (representing $l_{\text{max}} \approx 500$) our assumption of a linear clustering starts to break down. Therefore, we did not consider these multipoles in our subsequent calculations.

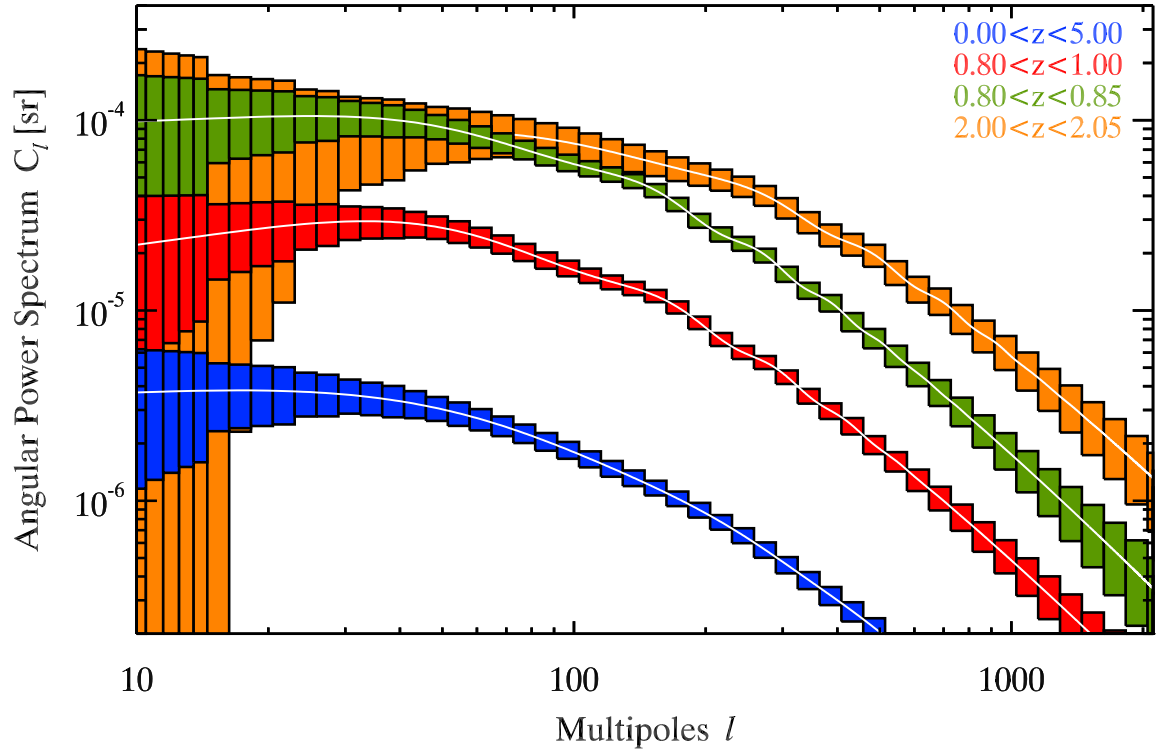


Figure 3.2: Same as Fig. 3.1, but with the angular power spectra for various narrow and broad redshift ranges added (see Sect. 3.2.2).

fixed effective mass of the DM halo (DMH). The details of our calculations are summarized below.

We calculated the angular power spectrum as follows:

$$C_\ell = \frac{2}{\pi} \int P(k) [W_\ell(k)]^2 k^2 dk \quad , \quad (3.1)$$

where the projection kernel is

$$W_\ell(k) = \int_{z_{\min}}^{z_{\max}} f(z) g(z) b(M_{\text{eff}}, z) j_\ell(k r(z)) dz \quad . \quad (3.2)$$

Here, $P(k)$ is the 3D linear power spectrum at $z = 0$, for which we used the fitting formulae of Eisenstein & Hu (1998), $f(z)$ is the normalized radial selection function, $g(z)$ is the linear growth function (e.g. Dodelson 2003), $b(M_{\text{eff}}, z)$ is the AGN linear clustering bias factor, and j_ℓ are the spherical Bessel functions on the order of ℓ , where $r(z)$ is the co-moving distance to redshift z (e.g. Hogg 1999).

The radial selection function is defined as the (normalized) differential redshift distribution of AGN, which we calculated with the AGN XLF, $\phi(\log L, z)$, of Hasinger et al. (2005) (also see Sect. 2.4 for details) It is the only quantity that contains the information about eRASS, since it depends on the survey sensitivity (S) as follows:

$$\frac{d\mathcal{N}}{dz}(S, z) = \frac{dV(z)}{dz} \int_{\log L_{\min}(S, z)}^{\log L_{\max}} \phi(\log L, z) d \log L \quad . \quad (3.3)$$

Here, $\frac{dV(z)}{dz}$ [Mpc³ sr⁻¹] is the co-moving volume element and $L_{\min}(S, z) = 4\pi S d_L^2(z)$, where $d_L(z)$ is the luminosity distance (e.g. Hogg 1999).

The AGN linear clustering bias factor, $b(M_{\text{eff}}, z)$, was computed with the analytical model of Sheth et al. (2001) by assuming an effective mass M_{eff} of the DMH where the AGN reside. Based on recent observations that cover the redshift range to $z \sim 3$ (e.g. Allevato et al. 2011; Krumpe et al. 2012; Mountrichas et al. 2013), we assumed an effective mass of $M_{\text{eff}} = 2 \times 10^{13} h^{-1} M_\odot$.

We only focused on the linear clustering regime. Therefore, we restricted our calculations to spatial co-moving scales larger than $k_{\max} \approx 0.2 h \text{ Mpc}^{-1}$, corresponding to wavelengths longer than $\approx 30 h^{-1} \text{ Mpc}$. The associated multipole number is $\ell_{\max}(\bar{z}) = k_{\max} r(\bar{z})$ and depends on the median redshift \bar{z} of the considered redshift bin. At the median redshift of eRASS AGN sample this is $\ell_{\max}(z \approx 1) \sim 500$. Thus, for our calculations we did not consider C_ℓ at multipoles higher than ℓ_{\max} .

For simplicity, we did not take into account linear redshift space distortions (RSD) (Kaiser 1987). Since the signal-to-noise ratio (S/N) in our angular power spectra is considerably poor at small multipoles (see Figs. 3.1 and 3.2) where the linear RSD become most significant, we do not expect that our results would change significantly if we would consider them in our calculations.

3.2.1 Uncertainties

The variance of the C_ℓ can be well approximated with

$$\sigma_{C_\ell}^2 = \frac{2}{(2\ell + 1) f_{\text{sky}}} \left(C_\ell + \frac{1}{\mathcal{N}} \right)^2 \quad (3.4)$$

assuming Gaussian statistics of the matter fluctuations ($\ell \lesssim \ell_{\text{max}}$). Here, f_{sky} is the sky fraction, which takes into account the effective loss of modes due to partial sky coverage, and \mathcal{N} is the AGN surface number density [sr^{-1}], which is computed with the AGN XLF and the survey sensitivity of eRASS (see Sect. 2.3 for details). The first term (C_ℓ) in the brackets represents the cosmic variance and becomes important at large scales (small ℓ). The second term, the shot noise (\mathcal{N}^{-1}), takes into account that we are using a discrete tracer (AGN) and becomes dominant at small scales (large ℓ), where $\mathcal{N}^{-1} \gg C_\ell$ (see e.g. Fig. 3.1). To minimize the uncertainty in C_ℓ , both a high sky coverage and a large number density of objects are needed.

3.2.2 Results

In Fig. 3.1, we show the expected angular power spectrum of the full eRASS AGN sample after four years for the entire extragalactic sky. By introducing redshift information (available from other surveys), the angular power spectrum becomes a relevant tool for LSS studies. In Fig. 3.2, we can see how its amplitude increases with a decreasing size of the redshift bin, and oscillations (see Sect. 3.4) in the angular power spectrum become more prominent as well. We can see from the two angular power spectra of $0.80 < z < 0.85$ and $2.00 < z < 2.05$ (with same redshift bin size) in Fig. 3.2 that the turnover of the spectrum and the positions of the oscillations depend on the redshift. The amplitudes are also different because the linear bias factor increases with redshift (see Sect. 3.3). Because the redshift distribution of eRASS peaks around $z \approx 0.8$, the number density around $z \approx 0.8$ is much higher than at $z \approx 2.0$ and therefore the uncertainty of the angular power spectrum is smaller for $0.80 < z < 0.85$ than for $2.00 < z < 2.05$.

3.3 Linear bias factor

The linear bias factor b is an important parameter for the clustering analysis of AGN. It connects the underlying DM distribution with the AGN population. Observationally, it has so far been a very challenging task to measure this connection with high accuracy, because of low statistics (e.g. Krumpe et al. 2010b, 2012; Miyaji et al. 2011; Starikova et al. 2011; Allevato et al. 2011, 2012; Mountrichas & Georgakakis 2012; Mountrichas et al. 2013). With more accurate observational knowledge of the behavior of the linear bias factor with redshift and luminosity and a comparison with simulations, we will be able to improve our understanding of major questions, such as the nature of the AGN environment, the main triggering mechanisms of AGN activity (e.g. Koutoulidis et al. 2013; Fanidakis et al. 2013) and how supermassive black holes (SMBH) co-evolve with the DMH over cosmic time (e.g. Alexander & Hickox 2012).

3.3.1 Method

The linear bias factor was measured by comparing the amplitudes of the observed power spectrum of tracer objects and of the theoretical power spectrum of the DM, under the assumptions of a certain cosmology. Since the power spectrum amplitude of tracer objects is proportional to the square of the linear bias factor ($\propto b^2$), its uncertainty directly reflects the uncertainty of measuring the latter. Knowing the amplitude (A) of our angular power spectrum, we are able to estimate this uncertainty. The S/N for measuring the normalization of the power spectrum C_ℓ assuming that its shape is known is given by

$$\frac{S}{N} = \frac{A}{\delta A} = \sqrt{\sum_{\ell=1}^{\ell_{\max}(\bar{z})} \left(\frac{C_\ell}{\delta C_\ell} \right)^2} \quad . \quad (3.5)$$

Here, we assumed that all multipoles are independent.

We used a redshift binning of $\Delta z = 0.20$ for our calculation (see Fig. 3.3 and 3.5) but other bin sizes would also be possible to demonstrate our results. In current observations (e.g. Allevato et al. 2011; Starikova et al. 2011; Koutoulidis et al. 2013) typically a much larger bin size is used to achieve a reasonable S/N for b in each redshift bin.

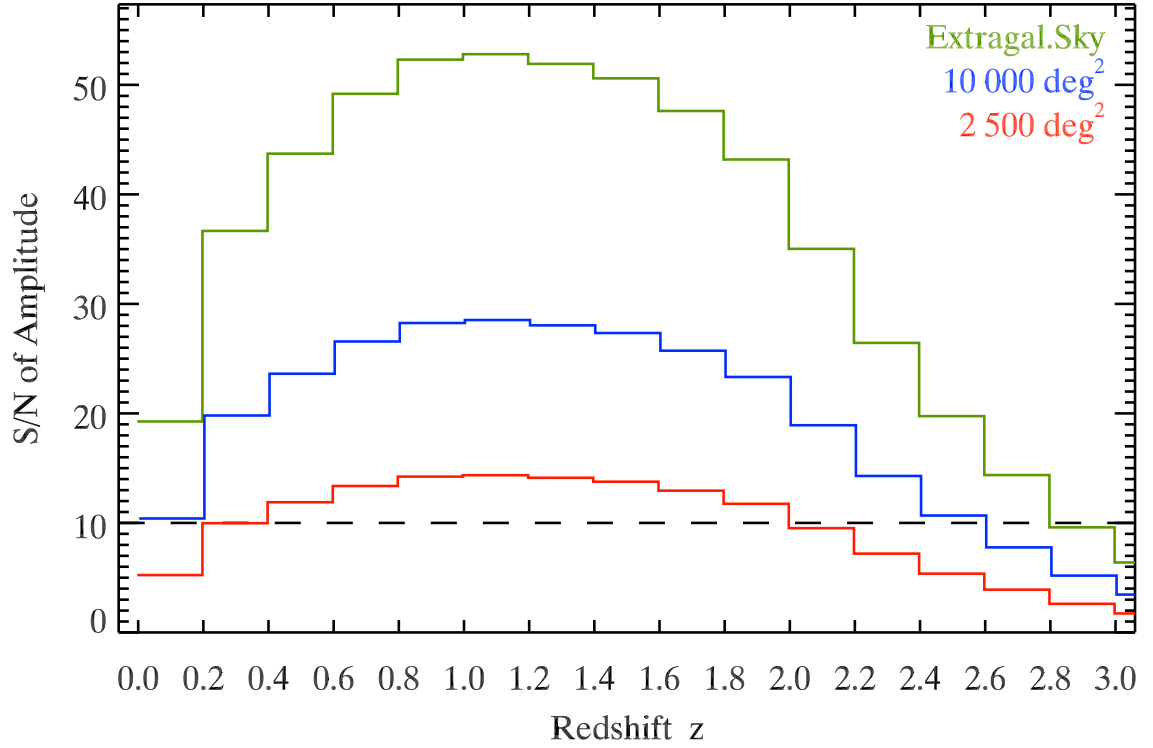


Figure 3.3: Signal-to-noise ratio of the amplitude of the angular power spectrum (Eq. 3.5) as a function of the redshift for different sky fractions. A $\Delta z = 0.2$ binning is assumed.

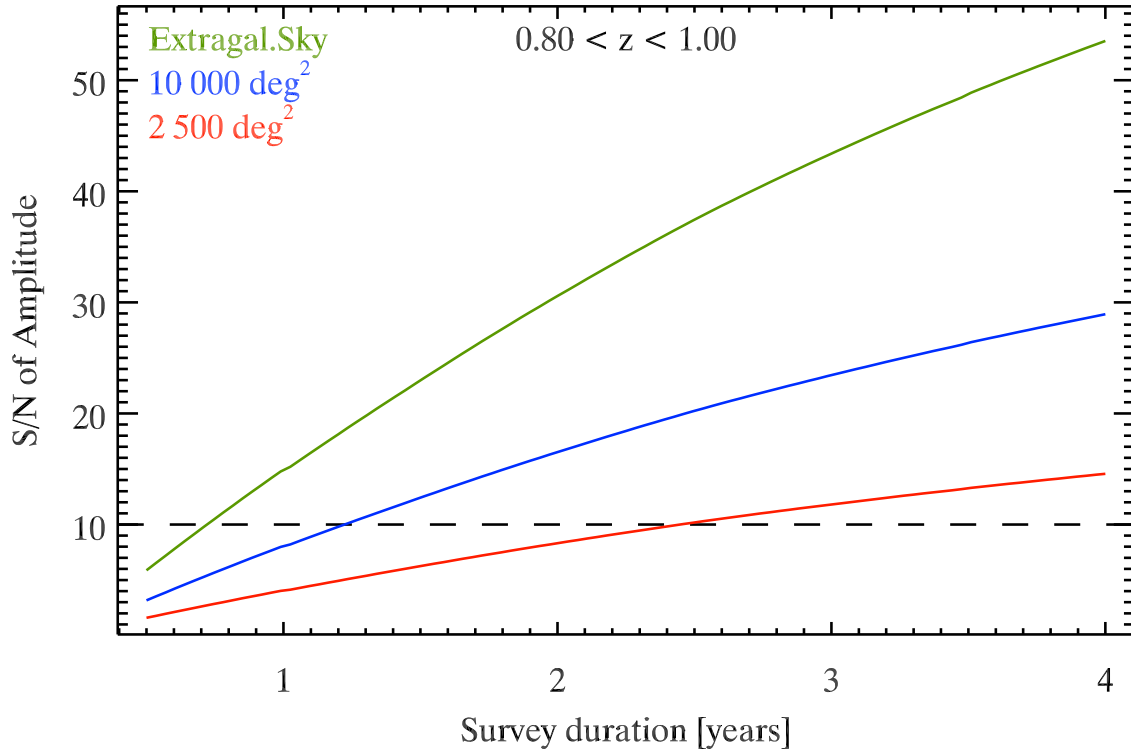


Figure 3.4: Signal-to-noise ratio of the amplitude of the angular power spectrum as a function of the survey duration for the redshift bin $0.8 < z < 1.0$ at different sky fractions.

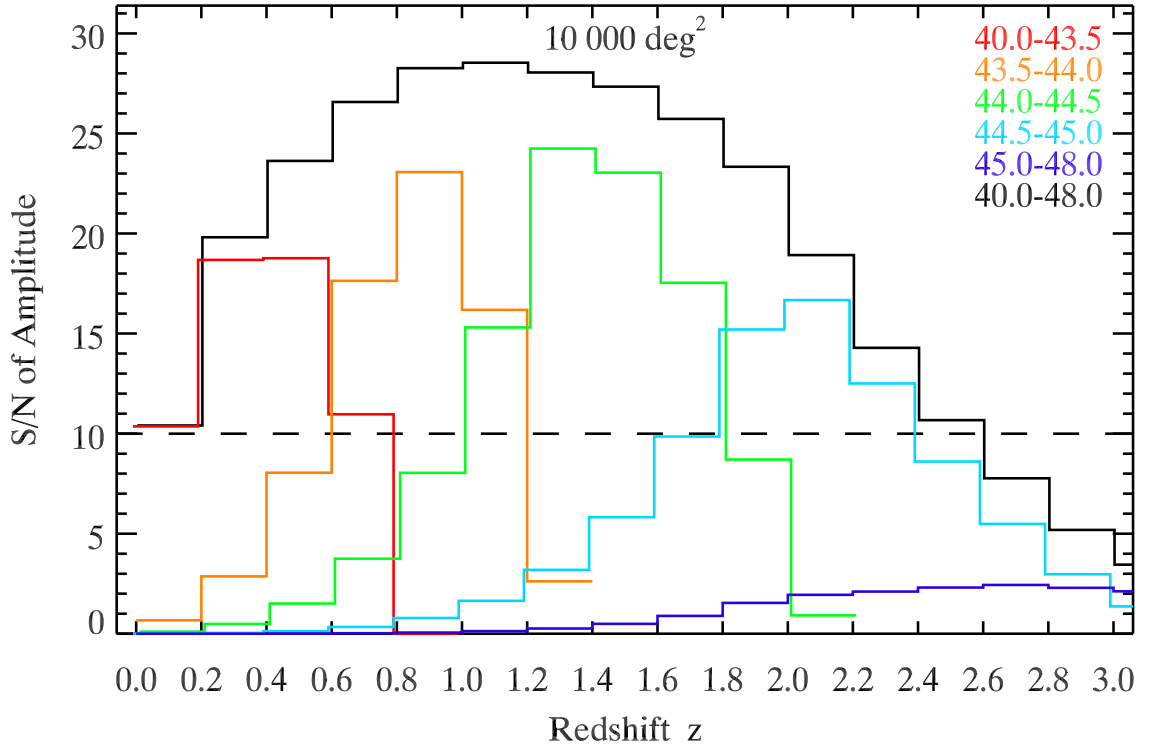


Figure 3.5: Same as Fig. 3.3, but for different luminosity ranges (in units of $\log(L[\text{erg s}^{-1}])$, 0.5 – 2.0 keV) and a sky coverage of 10 000 deg².

3.3.2 Results

In Fig. 3.3 the achievable S/N of the power spectrum amplitude is shown as a function of redshift. The shape of the curves is dominated by the redshift distribution of AGN modified by the quadratic-like increase of the linear bias factor with redshift at constant DMH mass (e.g. Sheth et al. 2001). We are able to measure the amplitude to a high accuracy ($< 10\%$) for a wide redshift range even with a fairly small fraction of the sky (e.g. $\sim 2500 \text{ deg}^2$).

The analysis of the linear bias factor can be performed before the entire four-year long eRASS is completed, as we can see from Fig. 3.4. For an SDSS-like sky coverage of 10000 deg^2 (blue curve) one can work with the data of eRASS after only 1.5 years (three full sky scans) to study the evolution of the linear bias factor to an accuracy of better than 10% in the amplitude for the redshift bin $0.8 < z < 1.0$. For a sky region of 2500 deg^2 it needs five full sky scans (2.5 years). For the neighboring redshift bins $0.6 < z < 0.8$ and $1.0 < z < 1.2$ the results are similar. The sensitivities used for this calculation are taken from Fig. 2.1.

Owing to the high S/N of the power spectrum amplitude, we will be able to separate the AGN into different luminosity groups. This is demonstrated in Fig. 3.5 for a sky coverage of 10000 deg^2 . We will be able to achieve an accuracy of $< 10\%$ for most luminosity groups for a certain redshift range. This means that it will be possible to perform a redshift- and luminosity-resolved analysis of the linear bias factor of AGN with eRASS with high statistical accuracy. We note that in our calculation the difference in the S/N of the luminosity groups in Fig. 3.5 is driven only by the difference in the redshift distribution of eRASS AGN and the redshift dependence of the linear bias factor.

3.4 Baryon acoustic oscillations

Acoustic peaks in the power spectra of matter and CMB radiation are among the main probes for measuring the kinematics of the Universe (e.g. Weinberg et al. 2013). They were predicted theoretically more than four decades ago (Sunyaev & Zeldovich 1970; Peebles & Yu 1970) and now have become a standard tool of observational cosmology. Unlike acoustic peaks in the angular power spectrum of CMB, their amplitude in the matter power spectrum in the Λ CDM Universe is small. For this reason, galaxy surveys have only recently reached sufficient breadth and depth for the first convincing detection of BAO, achieved with the SDSS data (Cole et al. 2005; Eisenstein et al. 2005; Hütsi 2006; Tegmark et al. 2006). Since then, BAO have been measured extensively up to redshift $z \sim 0.8$, in particular with luminous red giant

galaxies (LRGs) (e.g. Anderson et al. 2012). Above this redshift limit, BAO features were only found in the correlation function of the transmitted flux fraction in the Lyman- α forest of high-redshift quasars (Busca et al. 2013; Slosar et al. 2013), but have not yet been directly detected in galaxy distribution.

For the as yet uncharted redshift range from $z \sim 0.8$ up to ~ 2.0 , AGN, quasars and emission-line galaxies (ELGs) are proposed to be the best tracers for measuring BAOs, however, currently existing surveys do not achieve the required statistics for a proper detection (Sawangwit et al. 2012; Comparat et al. 2013). eRASS and the proposed SDSS-IV⁴ survey program eBOSS⁵ (2014-2020) will be the first surveys to change this situation. eRASS will achieve a sufficiently high density of objects $\mathcal{N} \sim 40 \text{ deg}^{-2}$ in this redshift range and will have by far the largest sky coverage compared with to eBOSS and all other dedicated BAO surveys. This would enable one to push the redshift limit of BAO detections in the power spectra of galaxies far beyond the present limit of $z \sim 0.8$.

3.4.1 Method

By construction of the model, BAOs are included in the AGN clustering model of Hütsi et al. (2012) through the 3D linear power spectrum (Sect. 3.2). Oscillations can be noticed, for example, in Fig. 3.2 in the power spectra of objects selected in narrow-redshift intervals. Because the angular scales of acoustic peaks depend on the redshift, BAO are smoothed out in the power spectra computed for broad-redshift intervals through the superposition of signals coming from many different redshift slices

Although the real data will be analyzed in a much more elaborate way, for the purpose of this calculation we used a simple method to estimate the amplitude and statistical significance of the BAO signal detection. We divided a broad redshift interval into narrow slices of width Δz and for each slice computed the angular power spectrum, $C_\ell(z)$, and converted the multipole number to the wavenumber $k = \ell/r(z)$ to obtain $P(k, z)$. These power spectra were co-added in the wavenumber space to obtain the total power spectrum $P(k)$ of objects in the broad-redshift interval. This power spectrum will have unsmeared BAO features. To estimate their statistical significance, we also constructed a model $C_{\ell, \text{smooth}}(z)$ without acoustic peaks by smoothing the matter transfer function, similar to Eisenstein & Hu (1998). From this model we computed the smoothed power spectrum in the wavenumber space, $P_{\text{smooth}}(k)$, that does not contain BAOs. To illustrate the amplitude of the BAO

⁴<http://www.sdss3.org/future/>

⁵http://lamwvs.oamp.fr/cosmowiki/Project_eBoss

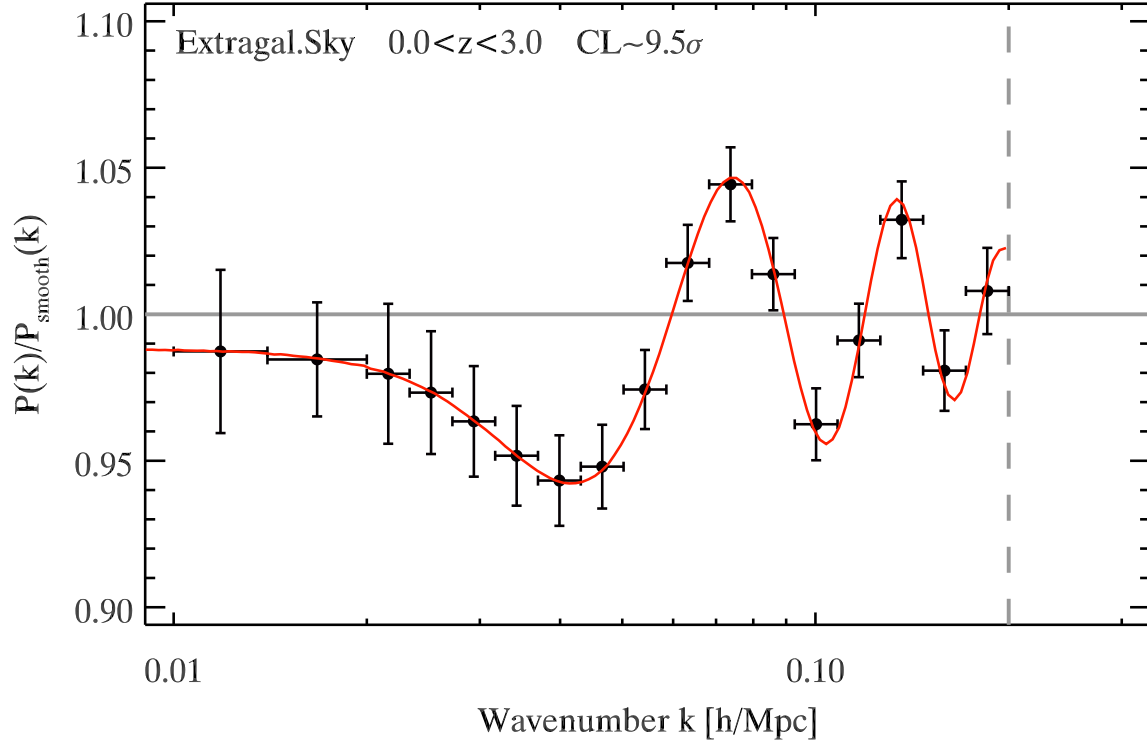


Figure 3.6: Baryon acoustic oscillations in the power spectrum for the extragalactic sky in the redshift range $0.0 < z < 3.0$. At wavenumbers above the vertical dashed line (corresponding to $0.2 h \text{ Mpc}^{-1}$) our assumption of a linear clustering starts to break down. The red curve shows the original input model for the BAOs.

signal, one can plot the difference $P(k) - P_{\text{smooth}}(k)$ or the ratio $P(k)/P_{\text{smooth}}(k)$.

By analogy with Eq. (3.5), the S/N of the BAO detection in the eRASS AGN sample can be computed as

$$\frac{S}{N} = \sqrt{\sum_z \sum_{\ell}^{\ell_{\text{max}}(z)} \left(\frac{C_{\ell}(z) - C_{\ell,\text{smooth}}(z)}{\sigma_{C_{\ell}}} \right)^2}, \quad (3.6)$$

where the outer summation was performed over the redshift slices and the variance $\sigma_{C_{\ell}}^2$ was calculated from Eq. (3.4).

The result of this calculation depends on the choice of the thickness of the redshift slice Δz (see Fig. 3.10). For too high values of Δz , BAO will be smeared out, as discussed above (cf. Fig. 3.2). On the other hand, for too low values of Δz , at which the thickness of the redshift slice becomes somewhat thinner than the co-moving linear scale of the acoustic oscillations, the cross-spectra between different redshift slices will need to be taken into account in computing $P(k)$. For the purpose of these calculations we chose $\Delta z = 0.05$. The corresponding thickness of the redshift slice at $z \sim 1$ is approximately equal to the co-moving linear scale of the first BAO peak. Note that omitting of the cross-spectra in our calculation leads to a slight underestimation of the confidence level of BAO detection.

3.4.2 Results

In Figs. 3.6 and 3.7 we show the ratio $P(k)/P_{\text{smooth}}(k)$ along with its uncertainties computed as described above. As these plots demonstrate, with the whole eRASS AGN sample for the extragalactic sky we expect to be able to detect the BAOs with a confidence level (CL) of $\sim 10\sigma$ (Fig. 3.6). For the currently unexplored redshift range of $0.8 - 2.0$ a confidence level of $\sim 8\sigma$ will be achieved, which can be seen in the top panel of Fig. 3.7. Decreasing the sky area to $20\,000 \text{ deg}^2$ or $10\,000 \text{ deg}^2$ (see middle panel of Fig. 3.7), we obtain $\sim 6\sigma$ and $\sim 4\sigma$, respectively. In Fig. 3.8 we show that the confidence level of the BAO detection for different redshift ranges depends on the sky coverage. The curves follow a $f_{\text{sky}}^{-0.5}$ -dependence, as expected from Eq. (3.4).

As Fig. 3.8 shows, for the redshift ranges $0.0 - 0.8$, $0.8 - 1.2$ and $1.2 - 2.0$ the achievable confidence levels are very similar, therefore the power spectra ratio shown in the bottom panel of Fig. 3.7 is representative for all three redshift intervals. Comparing the upper and bottom panels of Fig. 3.7, one can see that the BAO signal depends on the redshift range, while the top and middle panels show the degradation due to the reduced sky coverage.

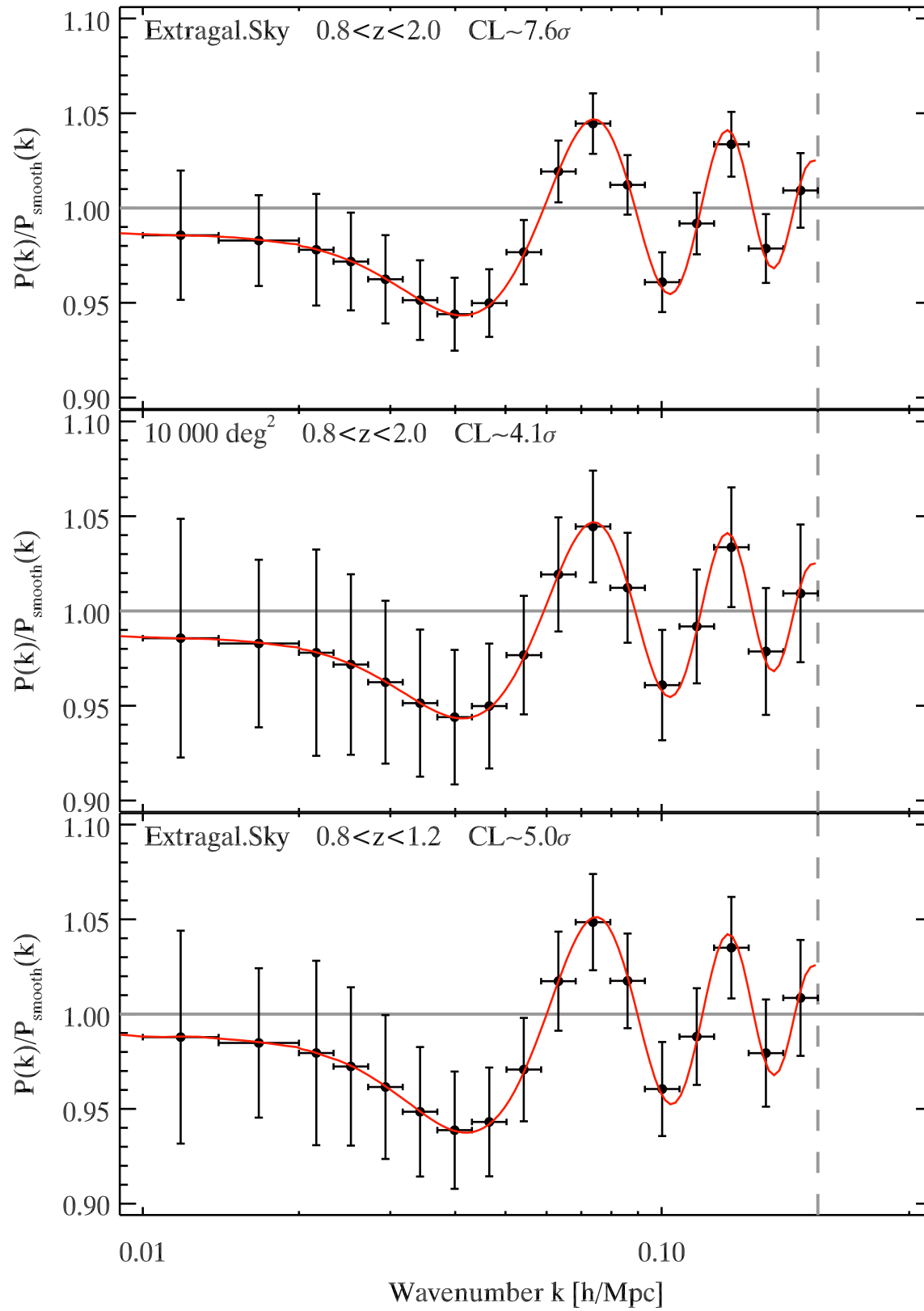


Figure 3.7: Same as Fig. 3.6, but for different redshift ranges and sky coverages.

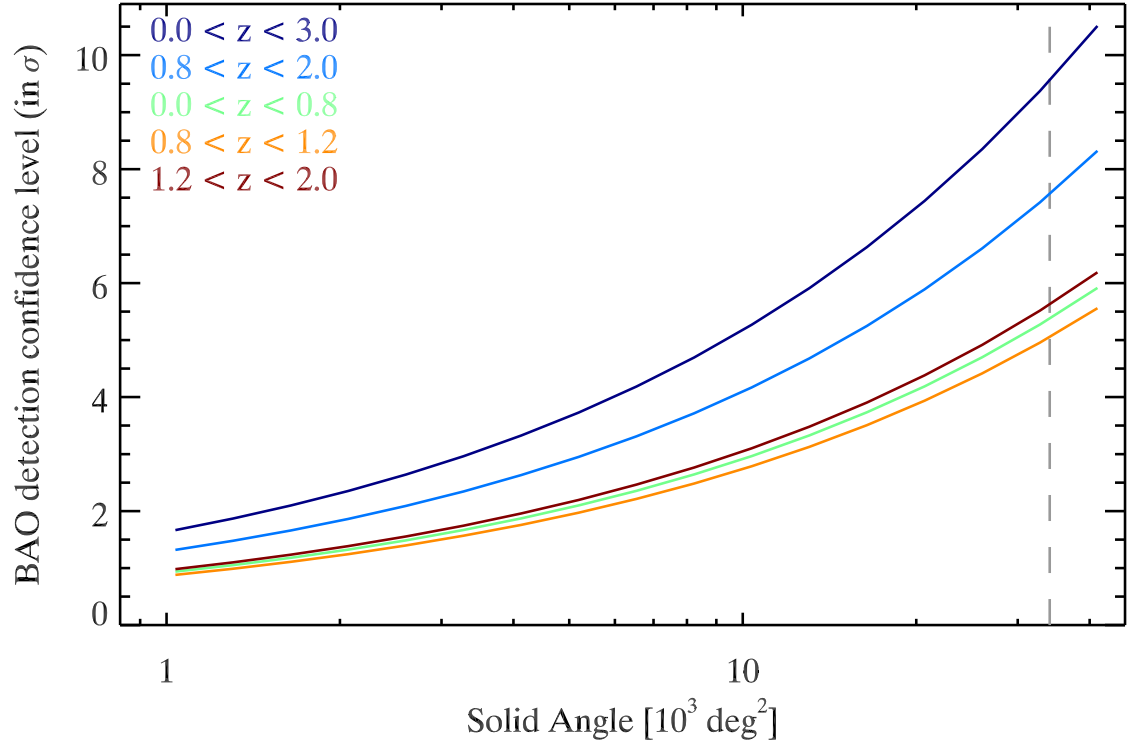


Figure 3.8: Confidence level of a BAO detection as a function of sky coverage for different redshift ranges (see Sect. 3.4.1 for more explanations). The vertical gray dashed line shows the area of the extragalactic sky.

3.5 Discussion and conclusions

3.5.1 Linear bias factor

Measuring of the linear bias factor provides a simple and direct method for estimating the average mass of DMHs that host a given subpopulation of AGN. With the eRASS, these measurements will become possible to unprecedented detail. The dramatic improvement of the redshift- and luminosity resolution of DMH mass measurements will have a great impact on our understanding of the environment of AGN, AGN triggering mechanisms, and SMBH co-evolution with the DMH.

Observational results of AGN clustering studies suggest a higher DMH mass for AGN than for quasars and a weak dependence between DMH mass and AGN luminosity (e.g. Krumpe et al. 2010b, 2012, 2014; Miyaji et al. 2011; Allevato et al. 2011, 2012; Cappelluti et al. 2012a; Mountrichas & Georgakakis 2012; Mountrichas et al. 2013; Koutoulidis et al. 2013; Fanidakis et al. 2013). However, uncertainties are still large and AGN luminosities available for these studies are typically limited by $L \sim 10^{44} \text{ erg s}^{-1}$. For instance, Koutoulidis et al. (2013) compared their results from clustering studies of AGN in four extragalactic X-ray surveys of different depth and coverage (CDFN, CDFS, COSMOS and AEGIS) with the theoretical predictions of Fanidakis et al. (2012). Their goal was to determine the dominant SMBH growth mode for AGN of different luminosities from the AGN bias factor measurements – either through galaxy mergers and/or disk instabilities or through accretion of hot gas from the galaxy halo. However, the uncertainties of bias factor measurements and of the DMH mass estimates were too large to clearly distinguish the dominant growth mode as a function of luminosity. In particular, the luminosity range of objects available for their analysis, $L \sim 10^{42-44} \text{ erg s}^{-1}$, was too narrow to challenge the prediction of Fanidakis et al. (2012) that luminous galaxies with $L > 10^{44} \text{ erg s}^{-1}$ reside in DMHs of moderate mass of $\sim 10^{12} M_{\odot}$. For the same reason, no direct comparison was possible with the results of the optical quasar surveys (Alexander & Hickox 2012). To overcome this limitation, Allevato et al. (2011) studied broad-line (BL) AGN from the COSMOS survey and found for their luminosity bin $L \sim 10^{43-46} \text{ erg s}^{-1}$, a significantly higher DMH mass than inferred from quasar studies, suggesting that for broad-line AGN major merger may not be the dominant triggering mechanism, which agrees reasonably well with recent simulations (Draper & Ballantyne 2012; Hirschmann et al. 2012; Fanidakis et al. 2012, 2013). However, the large width of the luminosity bin required to accumulate sufficient statistics did not allow them to draw a firm conclusion.

As illustrated by Fig. 3.5, the eRASS AGN sample will not only dramatically im-

prove the statistics, but will also expand the luminosity range beyond $L \sim 10^{44} \text{ erg s}^{-1}$ to the luminosity domain characteristic of quasars. Thus, the eRASS data will not only increase the redshift- and luminosity resolution of DMH mass estimate of AGN, but will open possibilities for a detailed comparison of the clustering properties of luminous AGN and optical quasars. Another aspect of bias measurements with eRASS that determines their uniqueness is that they are based on the X-ray selected AGN sample and will cover a very broad SMBH mass range, broader than that in AGN samples produced by optical/IR or radio surveys (e.g. Hickox et al. 2009).

The growth rate of SMBHs over time can be measured from the XLF of AGN (e.g. Aird et al. 2010) and eRASS will improve the accuracy and redshift resolution of these studies tremendously (Sect. 2.5). Combined with clustering bias data, these measurements will be placed in a broader context and be connected with DMH properties, which will provide new insights into the co-evolution of SMBHs with their DMHs (e.g. Alexander & Hickox 2012) and will also help to investigate the dominant triggering mechanisms of AGN activity (e.g. Koutoulidis et al. 2013; Fanidakis et al. 2013).

The AGN clustering model used here and the corresponding calculations of the AGN linear bias factor ignored the internal structure of DMHs, that is they were restricted to scales larger than the size of a typical DMH. Expressed in the language of the halo occupation distribution (HOD) formalism, these calculations operated with population-averaged halo occupation numbers. The angular resolution of the eROSITA telescope, $\approx 30 \text{ arcsec}$ FOV averaged HEW, is sufficient to resolve subhalo linear scales. Clustering measurements on small scales will permit one to obtain a detailed picture of the way AGN are distributed within a DMH (e.g. to measure fractions of central and satellite AGN), as well as how the HOD depend on the DMH mass and redshift, and AGN luminosity (Miyaji et al. 2011; Allevato et al. 2011; Starikova et al. 2011; Krumpe et al. 2012). Extrapolating the results of XMM-COSMOS data analysis by Richardson et al. (2013), we may expect that high-accuracy determination of the HOD parameters will be easily achieved with eRASS data, which will be able to address all these questions, advancing our understanding of AGN clustering on small scales and their HOD.

3.5.2 BAOs

The BAO detection beyond redshift ~ 0.8 will be a very significant milestone for the direct measurement of the kinematics of the Universe. eRASS will be able to map this uncharted redshift region up to $z \sim 2$ with a sufficiently high AGN number density to measure BAOs with a high statistical significance (see Fig. 3.8). For a proper

prediction of the way in which these measurements will improve our constraints on cosmological parameters, Markov chain Monte Carlo (MCMC) simulations (e.g. Lewis & Bridle 2002) and/or Fisher matrix calculations (e.g. Tegmark et al. 1997) have to be made, which is beyond the focus of our work. Sawangwit et al. (2012) have performed an MCMC simulation for quasars/quasi-stellar objects (QSOs) and demonstrated that a $3 - 4\sigma$ BAO detection (of a 3000 deg^2 QSO survey with $\mathcal{N} = 80 \text{ deg}^{-2}$) for $1.0 < z < 2.2$ can significantly reduce the uncertainties. Although the survey parameters of eRASS differ (much larger sky coverage but smaller source density for the same redshift region, $\sim 40 \text{ deg}^{-2}$), the results of Sawangwit et al. (2012) can give one an idea of how the eRASS AGN sample will improve the accuracy of cosmological parameter determination.

Our calculations were limited to the linear regime and did not take nonlinear structure growth into account, which would smear out the BAO signal to some extent. This would lead to a decreased detection significance (e.g. Eisenstein et al. 2007). However, with BAO reconstruction methods one will be able to correct for this effect to some extent (e.g. Padmanabhan et al. 2012; Anderson et al. 2012). We also note that our confidence level estimates of the BAO detection are fairly conservative because they neglect information contained in the cross-spectra. This will counterbalance the negative effect of BAO smearing, as we demonstrate in our subsequent study (Hütsi et al. 2014), where we study BAO predictions in a broader context and account for these effects more accurately.

3.5.2.1 Comparison with dedicated BAO surveys

We now compare the potential of the eRASS AGN sample with dedicated BAO surveys in the optical band. For the latter, we consider the completed BOSS CMASS survey (Anderson et al. 2012), the planned eBOSS⁶ and HETDEX (Hill et al. 2004) surveys that are scheduled to start in 2014, and the future BigBOSS survey (Schlegel et al. 2011), anticipated to be operational in 2020 time. Table 3.1 summarizes the key parameters of these surveys that are relevant for the BAO studies.

A quantity often used to estimate the statistical performance of a galaxy clustering survey is its effective volume (e.g. Eisenstein et al. 2005):

$$V_{\text{eff}}(k) = \Omega_{\text{survey}} \int_{z_{\text{min}}}^{z_{\text{max}}} \left[\frac{n(z) P_{\text{tr}}(k, z)}{n(z) P_{\text{tr}}(k, z) + 1} \right]^2 \frac{dV(z)}{dz d\Omega} dz, \quad (3.7)$$

⁶http://lamwms.oamp.fr/cosmowiki/Project_eBoss

Table 3.1: Parameters of BAO surveys.

Survey	Tracer Object	b_{tr} ($z = 0$)	Redshift Range	Ω_{survey} [10^3 deg^2]	\mathcal{N} [deg^{-2}]	$\langle n \rangle$ [$10^{-4} \text{ h}^3 \text{ Mpc}^{-3}$]	V_{eff} [$\text{h}^{-3} \text{ Gpc}^3$] ($k = 0.07 h \text{ Mpc}^{-1}$)	Implem. Date	Ref.
eRASS	AGN	1.33	0.8–2.0	34.1	40	0.12	7.8	2016–2020	
BOSS	LRG	2.00	0.4–0.7	10.0	80	2.3	2.9	concluded	[1]
eBOSS	ELG	1.00	0.6–1.0	1.5	180	2.1	0.5	2014–2020	[2]
eBOSS	QSO	1.20	1.0–2.2	7.5	90	0.21	1.7	2014–2020	[2]
BigBOSS	ELG	0.84	0.7–1.7	14.0	1730	6.3	24.0	> 2020	[3]
HETDEX	L α E	2.34 ^(a)	1.9–3.5	0.45	3370	7.2	1.9	2014–2017	[4]

^(a) for $z \approx 2.2$

Tracer object: AGN - active galactic nuclei, LRG - luminous red giant galaxy, ELG - emission line galaxy, QSO - quasar/quasi-stellar object, L α E - Lyman- α emitting galaxy

Columns: b_{tr} - bias factor of the tracer object; Ω_{survey} - solid angle covered by the survey; \mathcal{N} - surface number density; $\langle n \rangle$ - average volume number density; V_{eff} - effective survey volume at the first BAO peak
References: [1] Anderson et al. (2012) and Dawson et al. (2013); [2] eBOSS team, priv. comm.; [3] Schlegel et al. (2011); [4] HETDEX team, priv. comm.

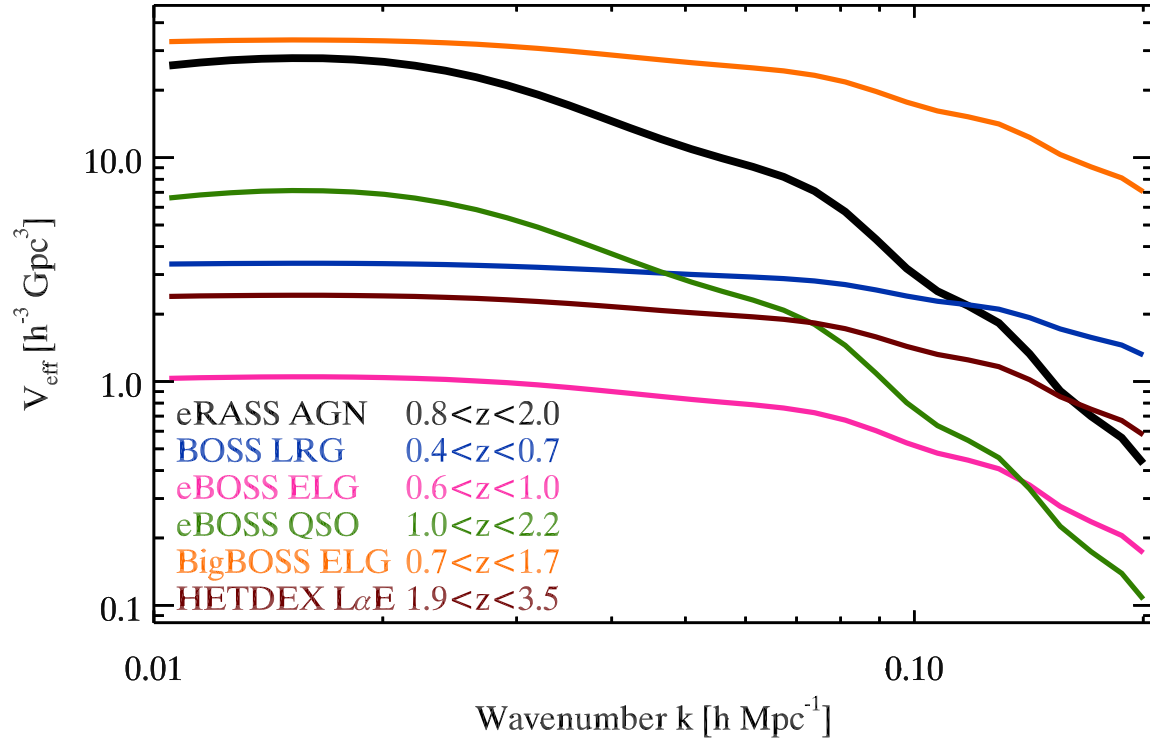


Figure 3.9: Effective volumes of BAO surveys listed in Table 3.1 as a function of the wavenumber. Effective volumes are computed for redshift ranges indicated in the plot.

where Ω_{survey} is the solid angle covered by the survey, $P_{\text{tr}}(k, z) = b_{\text{tr}}^2(z)[g(z)/g(0)]^2 P(k)$ is the power spectrum of objects used as LSS tracer, $b_{\text{tr}}(z)$ is their redshift-dependent bias factor and $n(z)$ is their redshift distribution [$\text{h}^3 \text{Mpc}^{-3}$]. Other quantities are defined in the context of Eqs. (3.1)–(3.3). For optical surveys we assumed that $b_{\text{tr}}(z)g(z) = \text{constant}$, therefore we only need to compute $P_{\text{tr}}(k, z = 0)$. For HETDEX we used $P_{\text{tr}}(k, z = 2.2)$, because the linear bias factor of HETDEX tracer objects was estimated by comparing the power spectrum of Lyman- α emitting galaxies (L α Es) from the simulations of Jeong & Komatsu (2009) at $z \approx 2.2$ with our linear (DM) power spectrum $P(k)$ transformed to $z = 2.2$. The $n(z)$ dependencies for optical surveys were taken from references listed in Table 3.1.

The results of these calculations are plotted in Fig. 3.9 where we show the effective volumes of different surveys as a function of the wavenumber. Their values at the first BAO peak are listed in the respective column of Table 3.1. In these calculations the integration in Eq. (3.7) was performed over the best-fit redshift range of each survey, as listed in Table 3.1. For eRASS we used the $z = 0.8 - 2.0$ range to emphasize its strength in this uncharted redshift region. The eRASS effective volume for the full redshift range is $\sim 70\%$ higher for the first BAO peak.

The result of the effective volume calculations obviously depends on the assumptions made of values and redshift dependences of co-moving density, bias, and the growth factor, which are not always precisely known, especially for the future surveys. Furthermore, efficiencies of redshift determinations are expected to be between 50 % and 80 %, but their exact values are difficult to predict. To have a fair comparison, we assumed a 100 % efficiency for all future surveys. Nevertheless, these curves should give a reasonably accurate comparison of the different surveys' qualities in measuring the power spectrum at different scales (note that the uncertainty of the power spectrum is proportional to $V_{\text{eff}}^{-0.5}$).

We can see from Fig. 3.9 that the effective volumes of eRASS AGN and eBOSS QSO samples fall more rapidly towards smaller scales than for other surveys. This is a consequence of the lower volume density of X-ray selected AGN and optical QSOs (Table 3.1). For the same reason, the statistical errors in the eRASS AGN power spectrum are dominated by the shot noise, but the high sky coverage of eRASS keeps them small. As one can see from the figure, eRASS is more competitive at larger scales, up to the second and third BAO peaks, where its sensitivity becomes similar to BOSS and HETDEX, respectively. It should be noted, however, that BOSS and HETDEX cover a relatively low ($z \lesssim 0.7$) and high ($z \gtrsim 2$) redshift domain, respectively, whereas all other surveys presented in Fig. 3.9 are aimed for the redshift region between ~ 0.8 and ~ 2.0 (Table 3.1). Around the first peak the effective volume of eRASS is a factor of about 2 – 4 higher than for BOSS and

HETDEX, but a few times lower than that of BigBOSS (Table 3.1). On the other hand, eRASS exceeds eBOSS at all wavenumbers. This would still be the case when one were to consider the subset of the eRASS sample to cover only $\sim 1/3$ of the extragalactic sky.

In conclusion we note that it is remarkable that the statistical strength of eRASS for BAO studies is similar to that of dedicated BAO surveys, even though eRASS was never designed for this purpose. Potentially, the eRASS AGN sample will become the best sample for BAO studies beyond redshift $z \gtrsim 0.8$ until the arrival of BigBOSS at the end of this decade. However, this potential will not be realized without comprehensive redshift measurements.

3.5.3 Redshift data

We assumed so far that the redshifts of all eRASS AGN are known. Now we briefly outline the requirements for the redshift data imposed by the science topics discussed above.

The linear bias as well as luminosity function studies do not demand a high accuracy of the redshift determination. Indeed, values of the order of $\delta z \sim 0.1 - 0.2$ should be sufficient, unless an analysis with a much higher redshift resolution is required. In principle, this accuracy can be provided by photometric surveys. However, one would need to investigate the impact of the large fraction of catastrophic errors, from which AGN redshift determinations based on the standard photometric filter sets are known to suffer (Salvato et al. 2011). Of particular importance are redshift- and luminosity trends in catastrophic errors. These problems are considered in our subsequent study (Hütsi et al. 2014). Provided that they are properly addressed, optical photometric surveys of a moderate depth of $I \gtrsim 22.5$ mag and with a sky coverage exceeding $\gtrsim 2500 \text{ deg}^2$ (Fig. 3.3) would already produce first significant results (see Sect. 2.7 for details). An existing survey with such parameters is SDSS. Its depth would allow detection of $\approx 80\%$ of eRASS AGN and with its sky coverage of $\sim 14500 \text{ deg}^2$ one should be able to conduct high-accuracy measurements of the linear bias factor. Of the other ongoing surveys, the Pan-STARRS PS1 3 π survey (Chambers & the Pan-STARRS team 2006) fulfills the necessary depth and sky coverage criteria.

BAO studies, on the other hand, require a much higher redshift accuracy of the order of $\delta z \sim 0.01$. This accuracy can be only achieved in spectroscopic surveys or in high-quality narrow-band multifilter photometric surveys. For example, for a 4σ detection of BAOs in the redshift range $0.8 < z < 1.2$, a spectroscopic survey of the depth of $I > 22.5$ mag is needed and sky coverage of at least $\sim 20000 \text{ deg}^2$

(Fig. 3.8). Promising candidates are the proposed 4MOST (de Jong et al. 2012) and WEAVE (Dalton et al. 2012) surveys, which would cover a large part of the sky with a multiobject spectrograph in the southern and northern hemisphere, respectively. An important caveat is that the angular resolution of eRASS (FOV averaged HEW of $\approx 30''$) is insufficient, in particular for faint X-ray sources, to provide accurate sky positions for spectroscopic follow-ups with multiobject spectrographs. Therefore additional photometric surveys (for example Pan-STARRS PS1 3π) will be needed to refine source locations to the required accuracy.

Fig. 3.10 shows the dependence of the confidence level of the BAO detection on the width of the redshift slice (Sect. 3.4.1). This plot roughly illustrates how the accuracy of the BAO measurement deteriorates with decreasing accuracy of the redshift measurements. Although the confidence level clearly degrades by a factor of $\sim 2 - 3$, the BAOs should still be (marginally) detectable even with fairly high values of $\Delta z \sim 0.2 - 0.3$, characteristic for errors of the photometric redshifts obtained with a standard set of broad photometric filters (Salvato et al. 2011). However, as we discussed earlier in this section, one of the significant problems of photometric redshifts is the large fraction of catastrophic errors. This factor was not accounted for Fig. 3.10. This problem is addressed in full detail in our subsequent study (Hütsi et al. 2014), which considers BAO data analysis in a more general context, including a realistic simulation of photometric redshift errors and cross-terms in the power spectra.

Finally, we note that we excluded from consideration a number of observational effects and factors, such as source confusion and source detection incompleteness, positional accuracy, telescope vignetting, and nonuniformities in the survey exposure, as well as several others. These factors and effects are well-known in X-ray astronomy, and data analysis methods and techniques exist to properly address them in the course of data reduction.

3.6 Summary

We have explored the potential of the eROSITA all-sky survey for large-scale structure studies and have shown that eRASS with its ~ 3 million AGN sample will supply us with outstanding opportunities for detailed LSS research. Our results are based on our previous study (Chapter 2), where we investigated statistical properties of AGN in eRASS, and on the AGN clustering model of Hütsi et al. (2012).

We demonstrated that the linear bias factor of AGN can be studied with eRASS to unprecedented accuracy and detail. Its redshift evolution can be investigated with an accuracy of better than $\sim 10\%$ using data from the sky patches of $\sim 2500 \text{ deg}^2$.

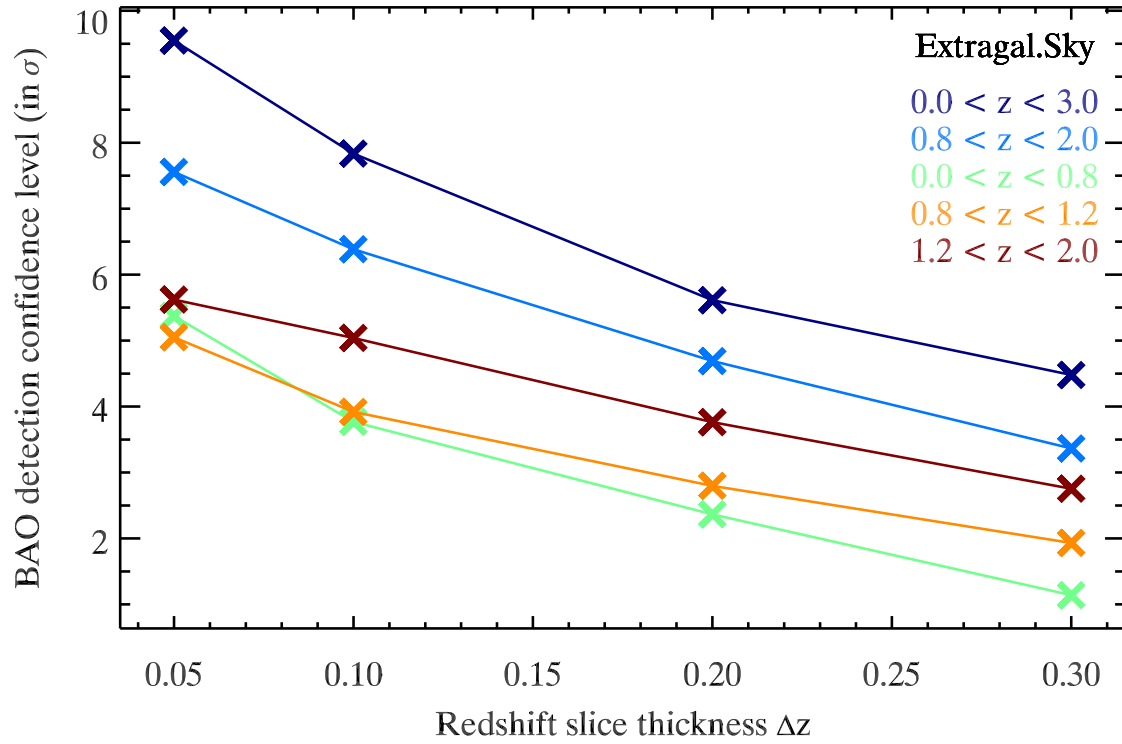


Figure 3.10: Confidence level of the BAO detection in the full extragalactic sky as a function of the redshift slice thickness (see Sect. 3.4.1) for different redshift ranges.

Using the data from a sky area of $\gtrsim 10\,000\,\text{deg}^2$, statistically accurate redshift- and luminosity-resolved studies will become possible for the first time. Bias factor studies will yield meaningful results long before the full four-year survey will be completed. The eRASS AGN sample will not only improve the redshift- and luminosity resolution of bias studies but will also expand their luminosity range beyond $L_{0.5-2.0\,\text{keV}} \sim 10^{44}\,\text{erg s}^{-1}$, thus enabling a direct comparison of the clustering properties of luminous X-ray AGN and optical quasars. These studies will dramatically improve our understanding of the AGN environment, triggering mechanisms, growth of supermassive black holes and their co-evolution with dark matter halos. The photometric redshift accuracy is expected to be sufficient for the bias factor studies, although the impact of the large fraction of catastrophic errors typical for standard broad-band filter sets needs yet to be investigated (Hütsi et al. 2014).

For the first time for X-ray selected AGN, eRASS will be able to detect BAO with high-statistical significance of $\sim 10\sigma$. Moreover, it will push the redshift limit of BAO detections far beyond the current limit of $z \sim 0.8$. The accuracy of the BAO investigation in this uncharted redshift range will exceed that to be achieved by eBOSS, which is planned in the same timeframe, and will only be superseded by BigBOSS, proposed for implementation after 2020. Until then, eRASS AGN can potentially become the best sample for BAO studies beyond $z \gtrsim 0.8$. However, for this potential to be realized and exploited, spectroscopic quality redshifts for large areas of the sky are required.

Part II

Large-scale structure studies with
the unresolved cosmic X-ray
background - challenges from
XBOOTES

Chapter 4

Surface brightness fluctuations in the unresolved cosmic X-ray background of XBOOTES

4.1 Introduction

Since the detection of the cosmic X-ray background (CXB) half a century ago (Giacconi et al. 1962), understanding of its origin has been one of the major drivers for X-ray astronomy and has led to the launch of most X-ray space telescopes up to the currently active missions NuSTAR, XMM-*Newton*, and *Chandra* (e.g. Fabian & Barcons 1992; Giacconi 2013; Tanaka 2013). Thanks to the many, in particular deep X-ray surveys of *Chandra* (e.g. Brandt & Hasinger 2005; Alexander et al. 2013), we now know for certain that the CXB is dominated by extragalactic point sources, with active galactic nuclei (AGN) leading the way (e.g. Comastri et al. 1995; Moretti et al. 2003; Hickox & Markevitch 2006, 2007; Gilli et al. 2007; Moretti et al. 2012). This makes the CXB the perfect window to study the accretion history of the Universe up to high redshift ($z \sim 5$) (e.g. Hasinger et al. 2005; Gilli et al. 2007; Aird et al. 2010; Ueda et al. 2014), which is an essential base to understand galaxy evolution (e.g. Hopkins et al. 2006; Hickox et al. 2009; Alexander & Hickox 2012).

Since the first X-ray surveys, studies of the brightness fluctuations had two major applications. They were used to disentangle the components of the CXB and at the same time to perform large-scale structure (LSS) studies (e.g. Barcons & Fabian 1988; Cappelluti et al. 2012a). With the high-angular resolution of the current generation of X-ray telescopes and complemented with spectroscopic redshift surveys of sufficient size the focus shifted towards clustering measurements of resolved AGN,

which developed in the last decade to an important branch of LSS studies. It led to major advances in understanding how AGN activity is triggered and how does it depend on its environment, such as the host galaxy and dark matter halo (DMH) properties, and how do supermassive black holes (SMBH) grow and co-evolve with their DMH over cosmic time (e.g Cappelluti et al. 2012a; Krumpe et al. 2014), which are essential questions in the field of galaxy evolution. In the future, it will be even possible to use AGN as a cosmological probe via baryon acoustic oscillation measurement (see Chapter 3 for details) with the ~ 3 million AGN to be detected by the upcoming eROSITA all-sky survey (see Chapter 2 for details).

Due the focus on resolved AGN, the current knowledge of AGN clustering properties and its implications for AGN and galaxy evolution are biased towards bright and luminous AGN ($L_{0.5-2.0\text{keV}} > 10^{42} \text{ erg s}^{-1}$), in particular for higher redshifts ($z > 0.5$), due to the signal-to-noise ratio (S/N) cut for the spectroscopic redshift and the luminosity cut from the AGN identification process (e.g Allevato et al. 2011, 2012, 2014; Krumpe et al. 2010b, 2012; Miyaji et al. 2011). An important question to ask is if we are able to extrapolate these clustering properties to less luminous AGN, which trace galaxies at an earlier evolutionary stage with a less massive SMBH and/or smaller accretion rate (Eddington ratio) than luminous AGN? A significant step towards answering this question is to study the brightness fluctuations of the unresolved CXB measured with the current generation of X-ray telescopes, which allows us to measure angular fluctuations on small scales down to the arc-second regime. The only two available studies of this kind analyzed the unresolved CXB with very deep surveys due to their scientific focus (e.g. Cappelluti et al. 2012b, 2013; Helgason et al. 2014). However, this also implies very small surveys in terms of sky coverage ($\sim 0.1 \text{ deg}^2$).

In our study we like to improve the current measurement of the brightness fluctuations of the unresolved CXB on small angular scales ($\lesssim 20'$) by an order of magnitude in S/N. We are able to achieve this by using the XBOOTES survey (Murray et al. 2005; Kenter et al. 2005, hereafter K05), the currently largest available continuous *Chandra* survey, with a surface area of $\sim 9 \text{ deg}^2$. The advantage in comparison to previous studies is that a higher S/N makes any discrepancies with current clustering models from known source populations more significant and it enables us to do clustering measurements in an energy resolved manner in order to separate different source populations. In this first study we present our measurement of the brightness fluctuations of the unresolved CXB with angular scales up to $\sim 17'$, and make novel tests for systematic uncertainties such as the brightness fluctuations of the instrumental background. Here, we analyze the average power spectrum of brightness fluctuations of all considered XBOOTES observations. This study is the essential

base of subsequent studies, where we will increase the complexity by analyzing a mosaic of all XBOOTES observations, which enables us to measure the fluctuations on angular scales up to $\sim 3^\circ$. With these studies we are able to cover the essential angular scales, where we can study the clustering properties of AGN within one DMH and AGN of different DMHs of low-luminosity AGN ($L_{0.5-2.0\text{ keV}} < 10^{42} \text{ erg s}^{-1}$) for redshift of $z > 0.5$. This parameter regime is inaccessible by clustering studies of the resolved CXB with current X-ray surveys (Cappelluti et al. 2012a).

Diffuse emission of the intracluster medium (ICM) of galaxy clusters & groups and the corresponding warm-hot intergalactic medium (WHIM) contributes significantly to the CXB (e.g. Hickox & Markevitch 2007; Roncarelli et al. 2012). Since galaxy clusters & groups are more difficult to detect and an order of magnitude more sparse than AGN, also our knowledge about their spatial distribution at low fluxes ($\lesssim 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$) is less certain (e.g. Rosati et al. 2002; Finoguenov et al. 2007, 2010; Clerc et al. 2012; Böhringer et al. 2014). Thanks to cosmological hydrodynamical simulations (e.g. Roncarelli et al. 2007, 2012; Ursino et al. 2011, 2014) and analytical studies (e.g. Diego et al. 2003; Cheng et al. 2004) we have nevertheless some reasonable understanding of their clustering properties. Here, clustering studies of the unresolved CXB could improved the situation from the observational side. However, at the depth of XBOOTES clustering measurements of galaxy clusters & groups are not very meaningful because the clustering signal of AGN dominates the brightness fluctuations of the unresolved CXB.

This study is organized as following: In Sect. 4.2 we explain our data processing procedure, in Sect. 4.3 we disentangle the different components of the unresolved CXB of XBOOTES, with focus on the unresolved AGN, and in Sect. 4.4 we present our measurement of the brightness fluctuations of the unresolved CXB, together with possible interpretations and tests for systematic uncertainties.

4.2 Data preparation and processing

For our analysis of the brightness fluctuations of the unresolved CXB we are using the XBOOTES survey (Murray et al. 2005, K05), which is currently the largest available continuous *Chandra* survey with a surface area of $\sim 9 \text{ deg}^2$. It consists of 126 individual, contiguous *Chandra* ACIS-I observations. In order to avoid unnecessary complication in our analysis, we exclude eight of them. The six observations with the ObsIDs 3601, 3607, 3617, 3625, 3641 & 3657 are excluded because they all show much higher background count rate than the average. The observations with the ObsIDs 4228 and 4224 are excluded because they contain the brightest point and extended source of the entire survey, respectively. Therefore, when referring to the

“XBOOTES survey”, we mean from now on the 118 remaining observations (see for a full list of observations Table 1 of Murray et al. 2005).

The average exposure time of a XBOOTES observations is ~ 5 ksec and the combined exposure time of 118 observations is almost 0.6 Msec. By excluding 8 observations the surface area of XBOOTES reduces from originally $\sim 9.3 \text{ deg}^2$ to $\sim 8.7 \text{ deg}^2$. Those values will further decrease after processing the observations.

For processing the observations of the XBOOTES survey we are using *Chandra*’s data analysis system CIAO (Scripts v4.6.7, CALDB v4.6.3, Fruscione et al. 2006) and follow their standard analysis threads, unless stated otherwise. Since the observations were performed in the very faint mode (VFAINT), we are able to make use of CIAO’s most strict filtering method¹ of background events in ACIS-I data.

Unless otherwise stated, we use throughout the paper for the Galactic absorption a hydrogen column density of $N_{\text{H}} = 10^{20} \text{ cm}^{-2}$ as determined for the XBOOTES survey and we convert the flux of extragalactic sources between different energy bands and between physical and instrumental units with an absorbed powerlaw with a photo-index of $\Gamma = 1.70$ (K05, Sect.3.3).

4.2.1 Exposure map and mask

For the following data processing and analysis, we will need the exposure map **E** [seconds] and mask **M**, which we describe here.

We use the exposure map **E** to convert our count maps **C** [counts] into flux maps **F** [cts s⁻¹] but also to take the vignetting into account. For creating the exposure map **E** we are using the same spectral model and parameters as for our spectral fit of the unresolved CXB (Table 4.4). We set all pixels of **E** to zero, where the exposure time falls below 63% of the peak value of **E** because we can see a clear break of the pixel distribution of **E** around this value. This method removes low exposed pixels, which are predominantly located in the CCD gaps and edges of the ACIS-I and occur due to the dithering movement of *Chandra* during an observation. Bad pixels and columns are automatically excluded with this process. The average field-of-view (FOV) solid angle of one observation after this filter step is $\sim 0.07 \text{ deg}^2$ (before removing resolved sources).

For the event-file of an observation, which will be used for the spectral analysis in Sect. 4.3, we can not use such a pixel filter. Instead, we reduce the FOV region of each ACIS-I CCD down to $\sim 92\%$ of the original size. This reduces the FOV almost in the same way at the CCD gaps and edges as with the 63% threshold of the exposure

¹ For details see http://cxc.harvard.edu/cal/Acis/Cal_prods/vfbkgrnd. We activate it in the data processing script `chandra_repro` with `check_vf_pha = yes`.

map. Further, we replace the exposure time in the header of the event-file with the average from the exposure map, which is more precise and leads to consistent flux values between the flux maps \mathbf{F} , used for the fluctuation analysis (Sect. 4.2.4-4.2.5), and the spectral fit of the unresolved CXB of XBOOTES (Sect. 4.3.2, Table 4.5).

We use the mask \mathbf{M} to excluded certain regions of our maps for our analysis. It has the same size as the exposure map, but pixels are either zero or one. Pixels area set only to zero, if they are outside the ACIS-I FOV, they have zero exposure time, or they are within the exclusion area of a resolved source (Sect. 4.2.2).

The mask and exposure map are set to be the same for all energy bands of an observation. For energy bands above 9 keV, where the *Chandra* has no effective area, we use an average exposure map

$$\langle \mathbf{E} \rangle = \mathbf{M} \cdot \left(\frac{\sum_i^N E_i}{\sum_i^N} \right), \quad (4.1)$$

when necessary and not otherwise stated. The summation goes over all pixels N , where the mask \mathbf{M} is one. The solid angle of the mask is computed as $\Omega = (\Delta p)^2 N$, whereby Δp is the size of a pixel. Since for our analysis we use an image pixel binning of one, the size of a pixel² is $\Delta p = 0.492''$.

4.2.2 Removing resolved sources

In order to study the unresolved CXB, we need to remove the resolved (point-like and extended) sources in the XBOOTES observations to such a level that the remaining counts of resolved sources contribute only insignificantly to the surface brightness of the unresolved CXB. For this purpose we do not perform our own source detection analysis. Instead we are using the two source catalogs of K05 for point and extended sources. This is also the reason why we need to remove the resolved sources first before we estimate and remove the instrumental background (Sect. 4.2.3 and 4.2.4).

4.2.2.1 Point sources

The point-source catalog includes 3293 sources with at least 4 counts in the 0.5 – 7.0 keV band. The sensitivity distribution for point sources as a function of sky coverage is shown in Fig. 12 of K05 (for the 0.5 – 7.0 keV band). The survey average sensitivity in the 0.5 – 2.0 keV band is $\sim 2.3 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$. To estimate the appropriate size of an circular exclusion area of a point source, we simulated the point-spread-function (PSF) shape for various offset angles (θ) from the aimpoint

²http://cxc.harvard.edu/proposer/POG/html/chap6.html#tab03:acis_char

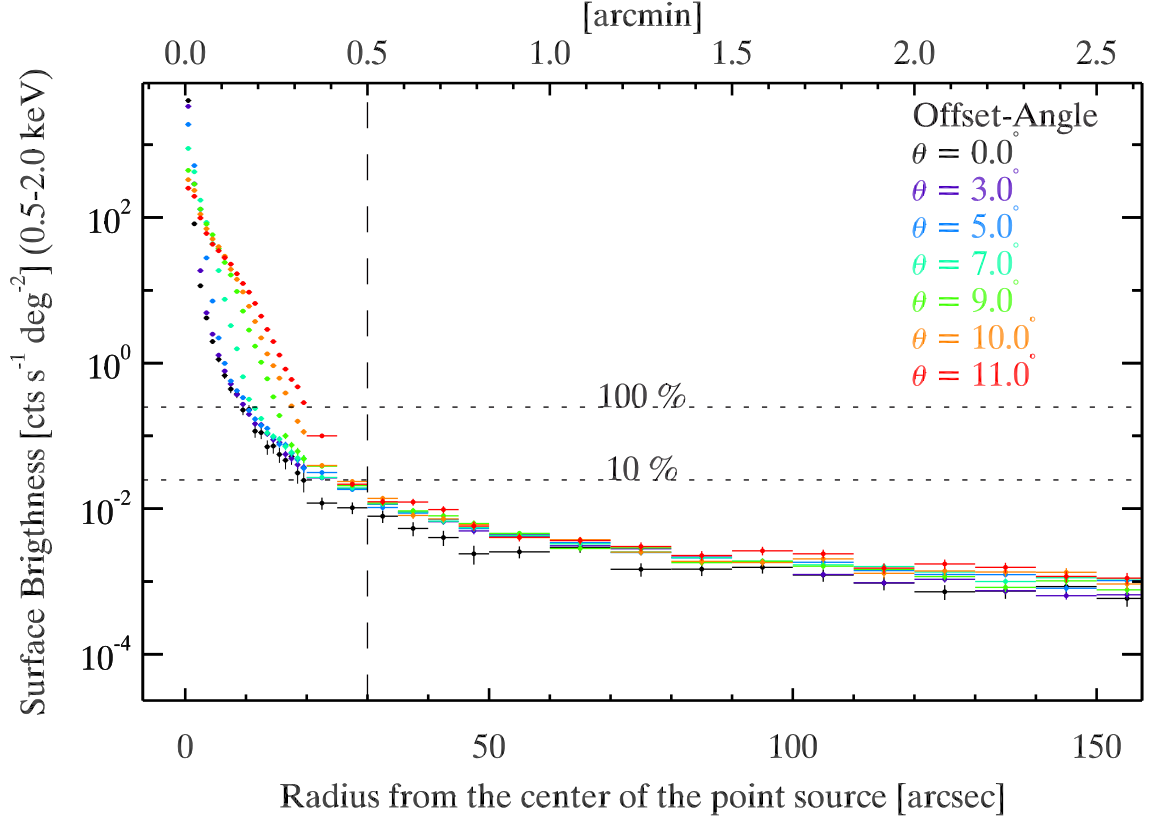


Figure 4.1: Shape of the PSF for *Chandra* ACIS-I (averaged over all CCDs) for different offset angles (θ) for the source flux of $0.63 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$ (0.5 – 2.0 keV). The vertical dashed line shows the radius of the circular exclusion area for this flux group (Table 4.1). The horizontal dotted lines show levels corresponding to 100 % and 10 % of the surface brightness of unresolved AGN (Table 4.7).

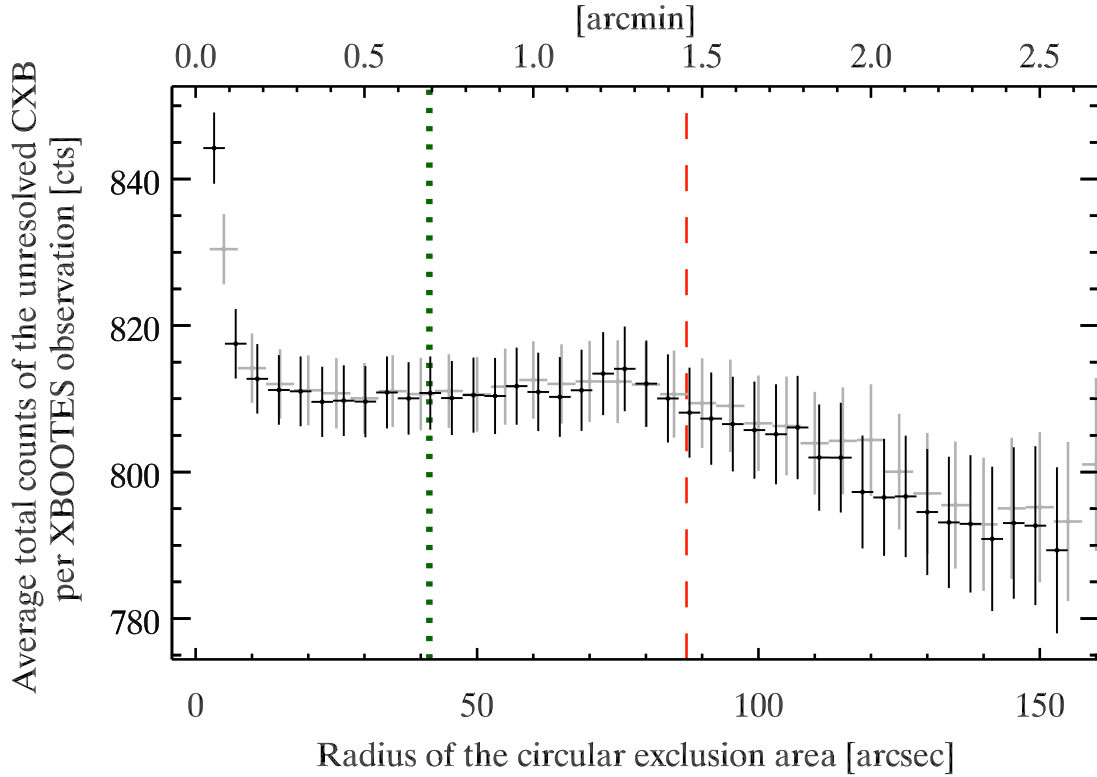


Figure 4.2: Average total counts of the unresolved CXB (0.5 – 2.0 keV) per XBOOTES observation as a function of radius of the circular exclusion area of point sources. Black crosses: using different radius intervals for the first three flux groups of Table 4.1 (see Text). The average radius was computed based on the number of point sources per flux group. Gray crosses: using the same radius interval for the first three flux groups. The radius of the fourth flux groups is fixed to the value in Table 4.1. Green dotted line: Average radius for the definition in Table 4.1. Red dashed line: radius, where the total removed area is equal to half the area of the ACIS-I FOV.

Table 4.1: Radius of the circular exclusion area of resolved point sources in XBOOTES for different flux groups.

	Flux Groups ^(a)	# of Sources	(b) [%]	Radius [arcsec]	(c) [%]
1)	[0.47, 0.63[1673	~ 51	30	~ 10
2)	[0.63, 2.10[1328	~ 40	55	~ 10
3)	[2.10, 9.00[268	~ 8	80	~ 25
4)	[9.00, 47.00[23	~ 1	220	~ 25

(a) In 10^{-14} erg cm $^{-2}$ s $^{-1}$ (0.5 – 2.0 keV).

(b) Number of sources in fraction of the total number of sources (3293).

(c) Upper limit of the point-source surface brightness at the edge of the exclusion area in fraction of the surface brightness of unresolved AGN (Table 4.7).

and for several ACIS-I CCD positions (ϕ) with the *Chandra* Ray Tracer simulator³ (Carter et al. 2003) and the MARX software package⁴ (v5.0.0), as shown in Fig. 4.1 for the average of all CCDs. Base on this simulation we define the circular exclusion area as presented in Table 4.1, where we split the point sources into different flux groups in order to make source removal more efficient. However, a separation into different offset angles or CCD positions is not required.

The observation with the ObsID 4228 contains the brightest point source of the entire XBOOTES survey with a flux of $\sim 1.6 \times 10^{-12}$ erg cm $^{-2}$ s $^{-1}$, which is about three times brighter than the second brightest point source. Due its large flux and the PSF shape with its extended wings (Fig. 4.1), it has the potential to contaminate the entire FOV. Therefore, we exclude the entire observation for our analysis (as mentioned in Sect. 4.2).

Evaluation In order to test that our definition of the exclusion area of point sources is sufficient, we estimated how the total counts of the unresolved CXB changes as a function of radius of the circular exclusion area (Fig. 4.2). To account that the PSF amplitude changes with flux, we use different radius intervals for the first three flux

³<http://cxc.harvard.edu/chart>

⁴<http://space.mit.edu/ASC/MARX>

groups of Table 4.1: 1) $[1'', 120'']$, 2) $[5'', 180'']$, 3) $[10'', 240'']$, in order to remove point sources to similar depth at a given evaluation step. Since this results in a different radius for each flux group at each evaluation step, we compute for presentation purposes an average radius weighted with number of point sources per flux group (Fig. 4.2). Since the fourth flux group contains only very few sources, we fix the radius for this flux group to the value in Table 4.1. In order to compare the total counts of the unresolved CXB between different evaluation steps, we correct them to the same surface area. For this we use the 9.5–12.0 keV band, where we can assume an homogeneous surface brightness for the entire FOV. To ensure a clean evaluation without any bias due to our choice of removing the extended sources (Sect. 4.2.2.2), we take for this test only those observations into account (83 out of 118 observations), which do not contain an extended source.

The result of the test is shown in Fig. 4.2 (black crosses). We can see for the radius interval $\sim 20'' - \sim 80''$ that the total counts of the unresolved CXB do not change significantly. The rise in total counts for $\lesssim 20''$ indicates that there is still a significant contamination by counts from resolved point sources. For large radii of $\gtrsim 80''$ we are removing more than half of the area of the FOV (red dashed line), where our surface area correction with the 9.5–12.0 keV band does not work properly anymore, which leads to a drop in total counts. The average radius of our definition of the exclusion area in Table 4.1 is $\approx 42''$ (weighted with number of point sources per flux group). We see that this radius (green dotted line) is located right in the middle of this interval. This proves that our definition is sufficient to remove the counts of resolved point sources but also is a good compromise in not removing too much surface area from the observation. In average the FOV is reduced by $\sim 17\%$ after removing all resolved point sources with our definition.

As an independent test we also used the same radius interval of $[5'', 200'']$ for the flux groups 1-3 of Table 4.1. This means, that they have the same radius for each evaluation step but also that point sources are removed at different depth. The result is also shown in Fig. 4.2 (gray crosses). We can see that total counts of the unresolved CXB followed the same trend as for our first test. Based on this we can conclude that in average for a *Chandra* ACIS-I observation at the depth of XBOOTES (~ 5 ksec), we are able to remove sufficiently well the counts of all resolved point sources with a flux of $F_{0.5-2.0 \text{ keV}} < 9 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ with a circular exclusion area of $\sim 20'' - \sim 80''$ in radius (independent of their offset angle).

4.2.2.2 Extended sources

There are 43 extended sources detected with a detection limit of $\approx 1 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ (0.5–2.0 keV) (K05, Sect. 3.2. & Table 1). However, the study of Vajgel et al. (2014) shows that only 32 of those are actually extended sources in the sense that there are more extended than the PSF at the same CCD position. The extended sources in the XBOOTES catalog were fitted with a Gaussian model in order to estimate their size. We define the radius of the circular exclusion area as six times this size. We tested circular exclusion areas between four and eight times the size and did not find any significant difference in the remaining source counts. Therefore, we believe that this is a reasonable definition. We also note that the total source counts of the resolved extended sources only accounts $\sim 4\%$ to the total source counts of all resolved sources, based on the source catalogs of XBOOTES.

The observation with the ObsID 4222 contains the brightest and widest extended source of the entire XBOOTES survey with a flux of $\sim 2.2 \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$ and a size of $\sim 44''$. It is about two times brighter than the second brightest and ~ 1.5 wider than the second widest extended source. Given this properties, this particular source has by far the highest potential to contaminate the entire observation even after removing most of its counts. Therefore, we exclude the entire observation for our analysis (as mentioned in Sect. 4.2).

4.2.2.3 Summary

After removing all resolved sources the average FOV area is reduced by $\sim 18\%$ down to $\sim 0.0610 \text{ deg}^2$. The average surface brightness is reduced by $\sim 43\%$ from 1.42 ± 0.01 to $0.81 \pm 0.01 \text{ ct s}^{-1} \text{ deg}^{-2}$ in the 0.5–2.0 keV band (after removing the instrumental background, see Sect. 4.2.3 and 4.2.4).

4.2.3 Removing background flares

In order to detect and remove time intervals of an observation, which are contaminated by background flares, we adopt the main concept of Hickox & Markevitch (2006, hereafter H06) and adjust them to the XBOOTES data. We analyze the light curve of each observation in the energy-band 2.3–7.3 keV. H06 show that this band is the best choice for background flare detection, because of the different energy-spectra of background flares and the quiescent background (see their Fig. 3).

Our de-flaring method consists of three consecutive steps of filtering the light curve: **(a)** We run a 3σ -clipping with the CIAO tool `deflare`, which is a standard procedure and removes the most obvious flares. Hereby, we use bins of $\sim 63 \text{ sec}$ (10

frames), which is large enough to assume a Gaussian error distribution in each time bin but small enough to not conceal short, strong flares.

(b) We create a light curve with a binning of ~ 252 sec (80 frames) and remove all bins, which are 30% above the mean count rate of the 3σ -clipped light curve from step (a). This step targets weaker and longer lasting flares with a maximum duration of the order of the bin size. In comparison to H06, we only remove positive deviations from the mean.

(c) We compute a light curve in bins of ~ 252 sec (80 frames) of the ratio between the 2.3 – 7.3 keV and 9.5 – 12.0 keV band and remove all bins, which are 40% above the mean ratio of all considered XBOOTES observations. This method was introduced by H06 and is best suited for weak flares. It takes advantage of the fact that for a typical flare the flux-ratio of 2.3 – 7.3 keV to 9.5 – 12.0 keV band will be larger than for the normal instrumental background (alias quiescent background) due to the different energy-spectrum shapes. We use the same threshold for all observations to ensure a constant energy-spectrum shape for all of them.

The major difference between H06 and our filtering arises due to the fact that our observations have exposure times of the order of kiloseconds, whereas H06 use observations with more than one Megasecond. This leads in our case to much smaller bin sizes for the light curves and less restrictive thresholds for removing flare events for step (b) and (c). The light curves of all observations were visually inspected and the thresholds of (b) and (c) were tuned to removed any obvious feature of the light curve, which could be interpret as a background flare.

For a typical observation, our de-flaring method removes in average ~ 190 sec ($\sim 4\%$). After the de-flaring we have an average exposure time per observation of ~ 4.3 ksec and a total exposure time is reduced to ~ 0.50 Msec.

Our observations might be still contaminated by weak flares. The brightness fluctuations of a flare is expected to follow a Poisson distribution. Therefore, even if an observation is still contaminated by a flares, we do not expect a significant distortion of the CXB brightness fluctuations.

4.2.4 Instrumental background and background-subtracted map

We estimate the contribution of the instrumental background with the method presented in H06. They show in their study with the *Chandra*'s ACIS-I stowed background data⁵ that the shape of the energy spectrum of the instrumental background

⁵<http://cxc.harvard.edu/contrib/maxim/acisbg/>

of ACIS-I from different observations is very stable over the course of five years, which includes the time when the XBOOTES observations were performed. Further, we know that all detected photons in the 9.5–12.0 keV band are due to the instrumental background because the effective area of *Chandra* in this energy range is neglectable. With those two facts combined we can estimate the *instrumental-background map* $\mathbf{C}_X^{\text{Obs.BKG}}$ of the *total-count map* $\mathbf{C}_X^{\text{Obs.}}$ in the energy band X by scaling the ACIS-I stowed-background map $\mathbf{C}_X^{\text{sto.BKG}}$ as following:

$$\mathbf{C}_X^{\text{Obs.BKG}} = \mathbf{M} \cdot \mathbf{C}_X^{\text{sto.BKG}} \cdot \left(\frac{\sum_i^N C_{9.5-12.0 \text{ keV}}^{\text{Obs.}}}{\sum_i^N C_{9.5-12.0 \text{ keV}}^{\text{sto.BKG}}} \right). \quad (4.2)$$

With this method we estimate an average background surface brightness of $1.55 \pm 0.01 \text{ cts s}^{-1} \text{ deg}^{-2}$ in the 0.5–2.0 keV band, which is consistent with the value from the spectral fit (see also Table 4.3). This means that $\sim 65\%$ of the total surface brightness of $2.37 \pm 0.01 \text{ cts s}^{-1} \text{ deg}^{-2}$ (after removing resolved sources, Sect. 4.2.2) is due to the instrumental background.

The *background-subtracted map* is then

$$\mathbf{C}_X^{\text{Scr.}} = \mathbf{C}_X^{\text{Obs.}} - \mathbf{C}_X^{\text{Obs.BKG}}. \quad (4.3)$$

We estimate for the background-subtracted map (after removing resolved sources) an average surface brightness of $0.81 \pm 0.01 \text{ cts s}^{-1} \text{ deg}^{-2}$ for the unresolved CXB of XBOOTES, which is in physical units $7.9 \pm 0.1 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$, using our spectral model of the unresolved CXB of XBOOTES (Sect. 4.3). As we can see in Table 4.5 this is consistent with the value from the spectral fit.

We use the instrumental-background estimation (Eq. 4.2) for measuring the surface brightness of the CXB (Table 4.5 and Sect. 4.3). However, for our fluctuation analysis (Sect. 4.4) we use for simplicity the total-count map $\mathbf{C}_X^{\text{Obs.}}$ instead of the background-subtracted map $\mathbf{C}_X^{\text{Scr.}}$. We can not detect any significant difference between those maps (see Sect. 4.4.4.2 for details) and for the total-count map we do not require any knowledge of the instrumental-background (also see Eq. 4.16).

4.2.5 Fluctuation maps

For our analysis we will need a *fluctuation map* $\delta \mathbf{F}$ in different energy bands for each observation. We compute this map for an energy band X as following:

$$\delta \mathbf{F}_X = \mathbf{F}_X - \langle \mathbf{F}_X \rangle. \quad (4.4)$$

Hereby, we compute the *flux map* \mathbf{F}_X from the ratio between the count and exposure map: $\mathbf{F}_X = \mathbf{C}_X / \mathbf{E}$. This definition takes the vignetting of the exposure map into

account, as long as we divide the count map with the proper exposure map \mathbf{E} and not with the average exposure map $\langle \mathbf{E} \rangle$ (Sect. 4.2.1). The effect of vignetting for our fluctuation analysis is discussed in Sect. 4.4.4.4. The average flux map $\langle \mathbf{F} \rangle$ is defined as:

$$\langle \mathbf{F}_X \rangle = \mathbf{M} \cdot \left(\frac{\sum_i^N C_X}{\sum_i^N E} \right) . \quad (4.5)$$

To compute the average count map $\langle \mathbf{C} \rangle$ we simply multiply $\langle \mathbf{F} \rangle$ with the exposure map: $\langle \mathbf{C}_X \rangle = \langle \mathbf{F}_X \rangle \cdot \mathbf{E}$.

4.3 Components of the unresolved CXB

The unresolved CXB⁶ consists of two components: the Galactic and extragalactic emission. With the help of the energy spectrum, we are able to separate their contributions to the CXB emission of the XBOOTES survey. We create this spectrum by stacking the energy spectra of all 118 considered XBOOTES observations.

The stacked energy spectrum has a total exposure time of ~ 0.50 Msec and is based on a total surface area of $\sim 7.2 \text{ deg}^2$ (without taking overlaps into account). We fit⁷ it in the energy range of $0.5 - 10.0 \text{ keV}$ with a source model (Sect. 4.3.2) and an instrumental background model (Sect. 4.3.1).

4.3.1 Instrumental-background model

We use the ACIS-I stowed-background map⁸ to create a spectral model for the instrumental background. We process this map as recommended by the CIAO threads. We fit this spectrum between 0.5 and 10.0 keV with a powerlaw for the continuum and with six Gaussians for the instrumental emission lines Al K_α , Si K_α , Au $K_{\alpha,\beta}$, Ni K_α and Au L_α , (Baganoff 1999, e.g. Fig. 3, left panel) and for an obvious feature around 8.3 keV , which is probably a complex of several weak emission lines. Since the spectral shape becomes more complex beyond 10.0 keV but does not improve our continuum estimation significantly, we limit the energy range to 10.0 keV .

The spectrum can be seen in Fig. 4.3. We can see that overall our model can well describe the energy spectrum of the instrumental background. However, we also

⁶We note that the term “CXB” is used ambiguously in the literature and some may use it exclusively for extragalactic emission.

⁷ With the X-Ray spectral fitting package XSPEC (v12.8.2, Arnaud 1996).

⁸<http://cxc.harvard.edu/contrib/maxim/acisbg/>; “acis[0-3]D2000-12-01bgstow_ctiN0004.fits”

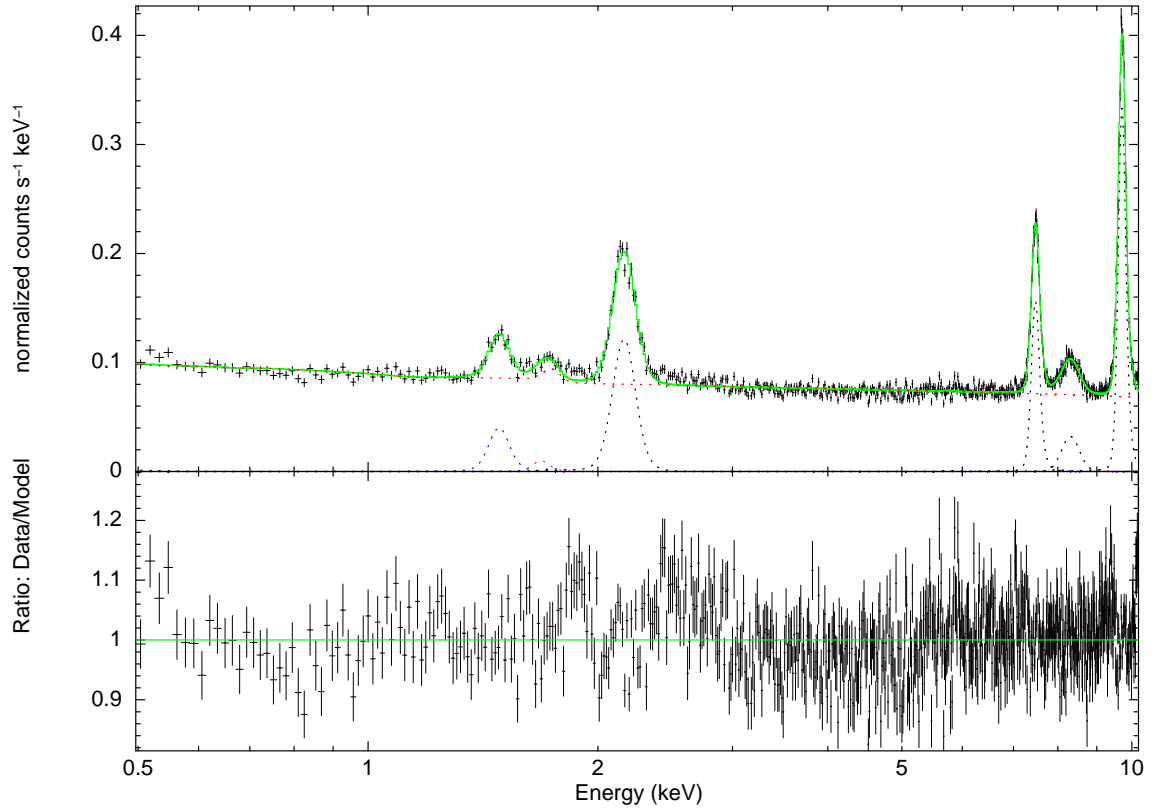


Figure 4.3: Energy spectrum of the ACIS-I stowed background to define our instrumental background model. Black crosses: data points with one standard deviation as error-bars. Green Curve: Total background model. Dotted Curves: single components of the model (see Table 4.2).

Table 4.2: Parameter values of the instrumental background model in the spectral fit of the unresolved CXB of XBOOTES (Sect. 4.3, Fig. 4.4-4.5)

Component-Name	Parameter	Value
powerlaw	Photon Index	0.107 (fixed)
1. Gaussian	Center	1.490 keV
(Al K_α)	Width	19.4 eV (fixed)
2. Gaussian	Center	1.678 keV (fixed)
(Si K_α)	Width	0.1 eV (fixed)
3. Gaussian	Center	2.162 keV
(Au $K_{\alpha,\beta}$)	Width	52.0 eV
4. Gaussian	Center	7.473 keV
(Ni K_α)	Width	11.8 eV
5. Gaussian	Center	9.700 keV
(Au L_α)	Width	35.0 eV
6. Gaussian	Center	8.294 keV
(Line complex)	Width	185.1 eV

The width is the standard deviation of the Gaussian. The fixed values are from the fit of the ACIS-I stowed background spectrum (Fig. 4.3).

Table 4.3: Surface brightness of the instrumental background of XBOOTES in the 0.5 – 2.0 keV band [cts s⁻¹ deg⁻²].

Estimated with	Value
Spectral fit (Sect. 4.3.1)	1.548 ± 0.004
Background-subtracted maps (Sect. 4.2.4)	1.55 ± 0.01

can see some deviations (see also the lower panel of Fig. 4.3). In Fig. 4.4 we can see that they are neglectable for $\lesssim 2.0$ keV because the source signal is much larger than those deviations. However, they become important for $\gtrsim 2.0$ keV, where the source signal is of the order of this deviations or smaller. Since we do not study the CXB brightness fluctuations for $\gtrsim 2.0$ keV, we do not see the need to improve our model in order to take those deviation into account.

All parameter values of the spectral fit of the of the unresolved CXB of XBOOTES are displayed in Table 4.2. For the this fit we keep all parameters free except the slope of the powerlaw, the width of the first Gaussian and the center and width of the second Gaussian. It is very difficult to obtain best-fit values for those parameters from this spectral fit and therefore we fix them to the values from the spectral fit of the ACIS-I stowed background (Fig. 4.3).

We can see in Fig. 4.4 that the background model has a higher continuum than the source model, which makes the estimate of the source flux sensitive to the powerlaw slope of the background model. It is reassuring that the estimate of the background surface brightness from the spectral fit is consistent with estimate from the background-subtracted maps (Sect. 4.2.4). The latter uses the method presented by H06 to estimate the background with the $9.5 - 12.0$ keV band (Eq. 4.2), whereas in our spectral fit we only fit up to 10.0 keV. This further supports that limiting the spectral fit to 10.0 keV does not negatively influence our instrumental background model.

4.3.2 Source model

The source model consists of an absorbed powerlaw (`phabs(powerlaw)`) with a fixed absorption column of $N_H = 10^{20} \text{ cm}^{-2}$ (K05) and with an unabsorbed `APEC`⁹ model. The spectrum and model fit are shown in Fig. 4.4-4.5 and the best-fit values of our source model are presented in Table 4.4. The source model gives a surface brightness of $7.8 \pm 0.1 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$ ($0.81 \pm 0.01 \text{ cts s}^{-1} \text{ deg}^{-2}$) in the $0.5 - 2.0$ keV band, which is in good agreement with the value from the flux maps (Sect. 4.2.4). All estimates of the surface brightness of the unresolved CXB of XBOOTES are summarized in Table 4.5. The individual components of our CXB model are discussed in the following sections.

We note that there is a significant hump in the energy spectrum around 2.5 keV

⁹ A collisionally-ionized diffuse gas model, based on the atomic database ATOMDB (v2.0.2), <http://www.atomdb.org>. Other diffuse gas models, such as `RAYMOND` or `MEKAL` are also appropriate. We use the solar abundances of Anders & Grevesse (1989), since it was used in several previous CXB studies.

Table 4.4: Best-fit parameters of our spectral model (APEC + phabs(powerlaw)) of the unresolved CXB of XBOOTES (Fig. 4.4-4.5).

Model-Component	Parameter	Value
APEC	Temperature (T)	0.164 ± 0.003 keV
	Normalization	$1.9 \pm 0.1 \times 10^{-4} \text{ cm}^{-5}$
	Surface Brightness ^(a)	$3.2 \pm 0.1 \times 10^{-12}$
phabs(powerlaw)	N_{H} (fixed)	10^{20} cm^{-2}
	Photon Index (Γ)	1.74 ± 0.03
	Normalization ^(b)	$1.23 \pm 0.02 \times 10^{-4}$
	Surface Brightness ^(a)	$4.6 \pm 0.1 \times 10^{-12}$

(a) $\text{erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$ (0.5 – 2.0 keV); (b) photons $\text{keV}^{-1} \text{ cm}^{-2} \text{ s}^{-1}$ at 1 keV.

Table 4.5: Surface brightness of the unresolved CXB of XBOOTES in the 0.5 – 2.0 keV band [$10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$].

Estimated with	Value
Spectral fit (Fig. 4.4, Sect. 4.3)	7.8 ± 0.1
Background-subtracted maps (Sect. 4.2.4)	7.9 ± 0.1

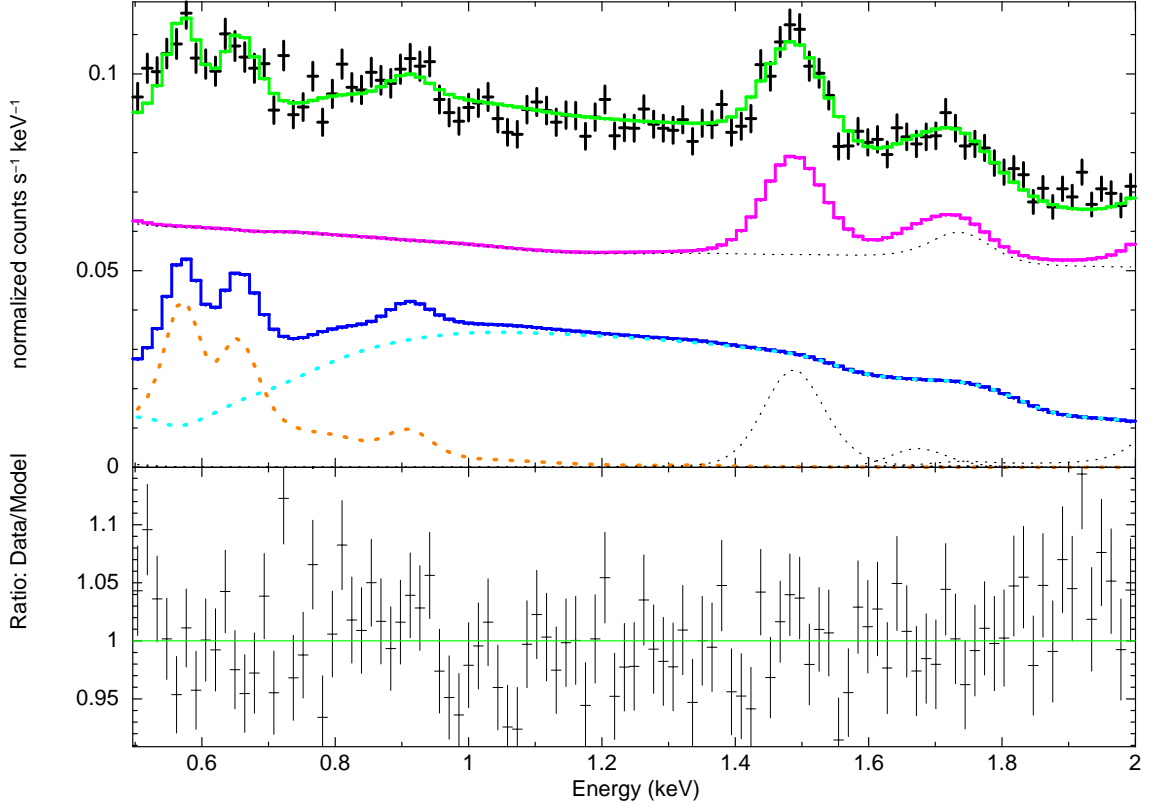


Figure 4.4: The stacked energy spectrum of the unresolved CXB from 118 XBOOTES observations (black crosses) for the 0.5 – 2.0 keV band, which we use for our brightness fluctuation analysis (Sect. 4.4), together with the ratio of model and data. Green solid curve: Total model (Source and background model). Dark blue solid curve: Source model (Table 4.4). Light blue dotted curve: absorbed powerlaw model (`phabs(powerlaw)`). Orange dotted curve: `APEC` (collisionally-ionized diffuse gas) model. Pink solid curve: Instrumental background model (Sect. 4.3.1). Black dotted curves: components of the instrumental background model.

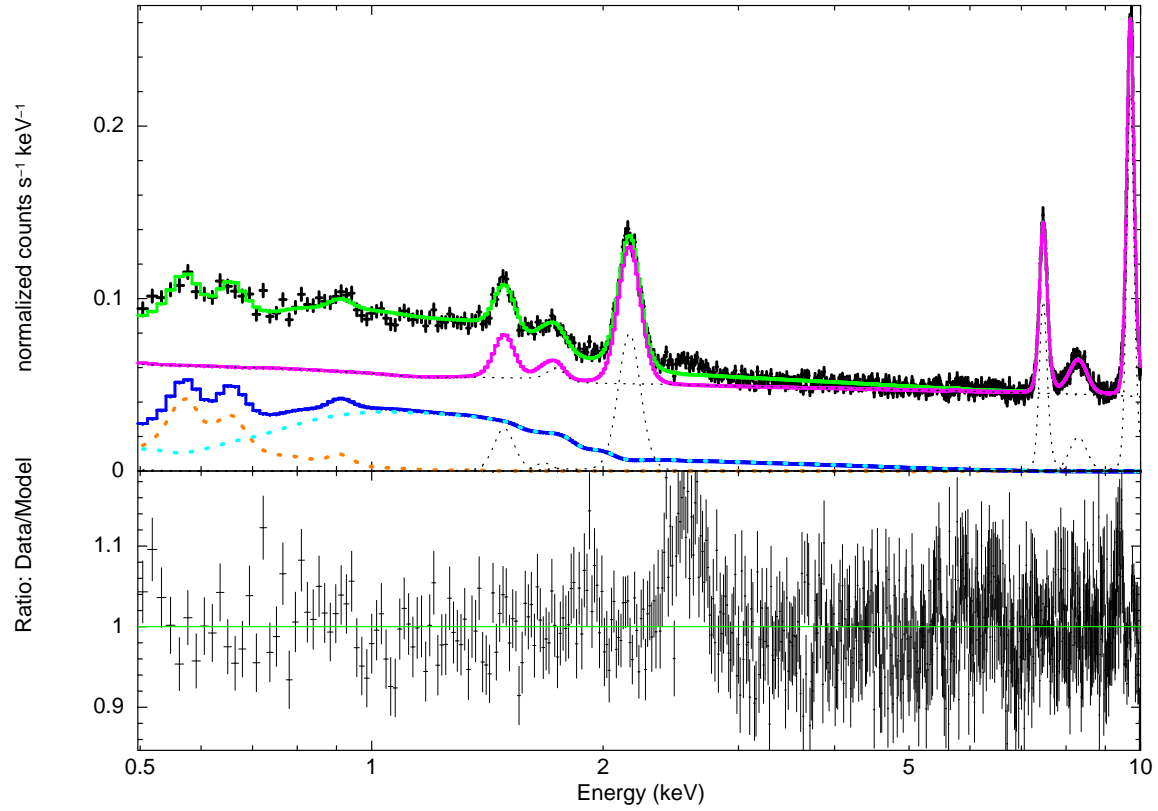


Figure 4.5: Same as Fig. 4.4 for the entire energy rang, which we use for the spectral fit.

(close to the right wing of the third instrumental line, Fig. 4.5). We believe that it arises from the instrumental background, since we can see a similar feature in the spectrum of the instrumental background (Fig. 4.3). This is further supported by the fact that any continuum emission from a real source would contribute much stronger at smaller energies than at 2.5 keV, because the effective area of *Chandra* ACIS-I is much larger there. However, we can not entirely exclude the possibility that this feature is a single emission line or line complex of a real source¹⁰. Since we focus in our fluctuation analysis (Sect. 4.4) in the 0.5 – 2.0 keV range, we do not investigate this particular spectral feature any further.

4.3.3 Galactic emission

The APEC model encapsulates all the Galactic emission, which is a superposition of various diffuse sources (e.g. Lumb et al. 2002; Hickox & Markevitch 2006; Henley & Shelton 2013): The Galactic halo emission and the foreground emission, which is a composite of emission from solar wind charge exchange (SWCX) and the local bubble. All of them have in common that they are anisotropically distributed over the sky. The Galactic emission dominates the soft part of the energy spectrum but is neglectable above ~ 1 keV, as we can see in Fig. 4.4.

The surface brightness of our Galactic emission model is $3.2 \pm 0.1 \times 10^{-12}$ erg cm⁻² s⁻¹ deg⁻² for the 0.5 – 2.0 keV band (Table 4.4). This is in reasonable agreement with the CXB measurements of H06 (Table 2, $\sim (3 - 4) \times 10^{-12}$ erg cm⁻² s⁻¹ deg⁻²), which use other *Chandra* surveys (*Chandra* Deep Fields, hereafter CDFs), and Lumb et al. (2002, Table 3, $\sim 3.8 \times 10^{-12}$ erg cm⁻² s⁻¹ deg⁻²), which use several deep XMM-*Newton* observations.

4.3.4 Extragalactic emission

The absorbed powerlaw describes the extragalactic emission and its total surface brightness is $4.6 \pm 0.1 \times 10^{-12}$ erg cm⁻² s⁻¹ deg⁻² for the 0.5 – 2.0 keV band. Together with the emission of the resolved sources ($4.4 \pm 0.1 \times 10^{-12}$ erg cm⁻² s⁻¹ deg⁻²) computed from the removed source counts and converted to physical units with the same spectral model, we obtain for XBOOTES a total extragalactic CXB surface brightness of $9.0 \pm 0.1 \times 10^{-12}$ erg cm⁻² s⁻¹ deg⁻². The values are summarized in Table 4.6. For the 1.0 – 2.0 keV band we have $5.1 \pm 0.1 \times 10^{-12}$ erg cm⁻² s⁻¹ deg⁻², which can

¹⁰ There is an example for such an emission line at ~ 3.5 keV, which may originate from the decay of sterile neutrinos (e.g. Bulbul et al. 2014), but its detection and interpretation is highly debated (e.g. Anderson et al. 2014).

Table 4.6: Extragalactic emission of XBOOTES [$10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$].

	0.5 – 2.0 keV	1.0 – 2.0 keV
Unresolved ^(a)	4.6 ± 0.1	2.62 ± 0.03
Resolved ^(b)	4.4 ± 0.1	2.46 ± 0.04
Total	9.0 ± 0.1	5.1 ± 0.1

(a) From the spectral fit (Table 4.4); (b) From the removed source counts of the resolved sources in the background-subtracted maps (Sect. 4.2.2-4.2.4).

be easily compared with the summary of H06 (Table 6, last column) of previous CXB studies. This comparison reveals that our measurement is in good agreement with these studies. The advantage of comparing the different measurement in the 1.0 – 2.0 keV instead of 0.5 – 2.0 keV band is, that we can ignore the contribution of the Galactic emission, which is $< 0.1 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$ for our spectral model.

The unresolved extragalactic emission of XBOOTES is a superposition of the emission of various unresolved sources (e.g. Lumb et al. 2002; De Luca & Molendi 2004; Hickox & Markevitch 2006, 2007; Kim et al. 2007; Georgakakis et al. 2008; Lehmer et al. 2012, hereafter L12): AGN, normal galaxies (no active nuclei), and galaxy clusters & groups. Its difficult to estimate the exact contribution of each component to the total unresolved emission (e.g. Hickox & Markevitch 2007). We can estimate the contribution of AGN and normal galaxies with the $\log N - \log S$ from the literature down to the flux limit of $\sim 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1}$ (e.g. Kim et al. 2007; Georgakakis et al. 2008, L12). Their surface brightness values are presented in Table 4.7. The $\log N - \log S$ of galaxy clusters & groups is not very well know below $\sim 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ (e.g. Rosati et al. 2002; Finoguenov et al. 2007, 2010; Clerc et al. 2012). Therefore, we give a range in Table 4.7, which shows the spread of different studies. The combined surface brightness of X-ray point sources and galaxy clusters & groups is $\sim (1.9 - 2.2) \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$, which accounts for about half of the measured extragalactic emission of the unresolved CXB of XBOOTES.

The remaining emission of $\sim (2.1 - 2.4) \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$ can not be easily disentangled. For a better comparison with previous studies, we use the value of $\sim (1.2 - 1.4) \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$ in the 1.0 – 2.0 keV band, where we can neglected Galactic emission. The currently deepest X-rays surveys (CDFs), which are used to compute the $\log N - \log S$ of X-ray point sources and galaxy clusters

Table 4.7: Individual components of the unresolved extragalactic emission of XBOOTES in the $0.5 - 2.0$ keV band [10^{-12} erg cm $^{-2}$ s $^{-1}$ deg $^{-2}$].

Extragalactic components	Surface Brightness	Ref.
Our measurement:	4.6 ± 0.1	Sect. 4.3.4
Estimations based on previous studies:		
AGN ^(a)	~ 1.5	[1]
Normal galaxies ^(a)	~ 0.2	[1]
Galaxy clusters ^(b)	$\sim (0.2 - 0.5)$	[2]-[3]
Unresolved emission ^(c)	$\sim (1.4 - 2.1)$	[4]
Sum of all components	$\sim (3.2 - 4.5)$	

(a) Using the sensitivity distribution of XBOOTES for point sources (Fig. 12 of K05, also see Sect. 4.2.2).

(b) Assuming a flux limit of 1×10^{-14} erg cm $^{-2}$ s $^{-1}$ ($0.5 - 2.0$ keV, K05, also see Sect. 4.2.2).

(c) Still unresolved in the currently deepest X-ray surveys (Hickox & Markevitch 2007). Converted from $1.0 - 2.0$ keV band and assuming a cosmic variance of 20 %. Ref.: [1] Lehmer et al. (2012); [2] Clerc et al. (2012); [3] Rosati et al. (2002); [4] Hickox & Markevitch (2006, Table 5);

& groups, still have an unresolved component of the CXB with an average surface brightness of $1.0 \pm 0.1 \times 10^{-12} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$ in the 1.0–2.0 keV band (H06). Note that this emission is not included in our calculations of contribution of AGN, normal galaxies, and galaxy clusters & groups with the $\log N - \log S$ from the literature. Since these are deep pencil-beam surveys, they suffer strongly from cosmic variance, which adds an additional uncertain of $\sim 20\%$ to this value (H06, Sect. 8.2.4.). In this respect, our remaining emission appears to be consistent with this measurement.

In summary, we can conclude that the surface brightness of the total and unresolved extragalactic emission of XBOOTES are fully consistent with previous CXB studies. This also confirms that we correctly estimate the instrumental background.

4.3.4.1 Unresolved AGN population

In our fluctuation study (Sect. 4.4.6) we will focus on the unresolved AGN. Therefore, we like to estimate some basic properties of them from their differential flux distribution (also known as the flux production rate per solid angle) as a function of redshift and luminosity:

$$\frac{dS(z)}{dz} = \int dS \left\{ \frac{\Phi(\log_{10}(L_r[S, z]), z)}{L_r[S, z] \log_e(10)} \right\} \frac{d^2V(z)}{dz d\Omega} \frac{4\pi d_L^2(z)}{K(z)} S [1 - f(S)], \quad (4.6)$$

$$\frac{dS(L)}{d \log_{10}(L)} = \int dz \Phi(\log_{10}(L_r), z) \frac{d^2V(z)}{dz d\Omega} S[L, z] [1 - f(S[L, z])]. \quad (4.7)$$

Here, $\Phi(\log_{10}(L_r), z) [\text{h}^3 \text{ Mpc}^{-3}]$ is the AGN X-ray luminosity function (XLF) of Hasinger et al. (2005), $f(S)$ is the survey selection function of XBOOTES for point sources (K05, Fig. 12), $K(z)$ is the K-Correction, $L_r [\text{erg s}^{-1}]$ is the rest-frame luminosity ($L_r = L/K(z)$), $d^2V(z)/dz/d\Omega [\text{Mpc}^3 \text{ h}^{-3} \text{ deg}^{-2}]$ is the co-moving volume element, and $d_L(z) [\text{cm}]$ is the luminosity distance (e.g. Hogg 1999). For the XLF we include the exponential redshift cutoff for $z > 2.7$: $\Phi(z) = \Phi(z=2.7) \times 10^{0.43(2.7-z)}$, which was proposed by Brusa et al. (2009). For the K-Correction we assume a powerlaw with the photo-index of $\Gamma = 1.7$, which simplifies the quantity to $K(z) = (1+z)^{2-\Gamma}$. The survey selection function is given in the 0.5 – 7.0 keV band and we convert it to the 0.5 – 2.0 keV band with an absorbed powerlaw with a photo-index of $\Gamma = 1.7$. Note, that the XLF of Hasinger et al. (2005) only includes type 1 AGN, which have a minimum luminosity of $L_{0.5-2.0 \text{ keV}} \approx 10^{42} \text{ erg s}^{-1}$ and maximum redshift of $z \approx 3 - 5$. We correct the amplitude of the differential flux distributions with the $\log N - \log S$ of L12 (factor of ~ 1.43 increase), but we can not correct the shape. Hence, one should be aware that these distributions may have high statistical uncertainties at low luminosity ($< 10^{42} \text{ erg s}^{-1}$) and very high redshift

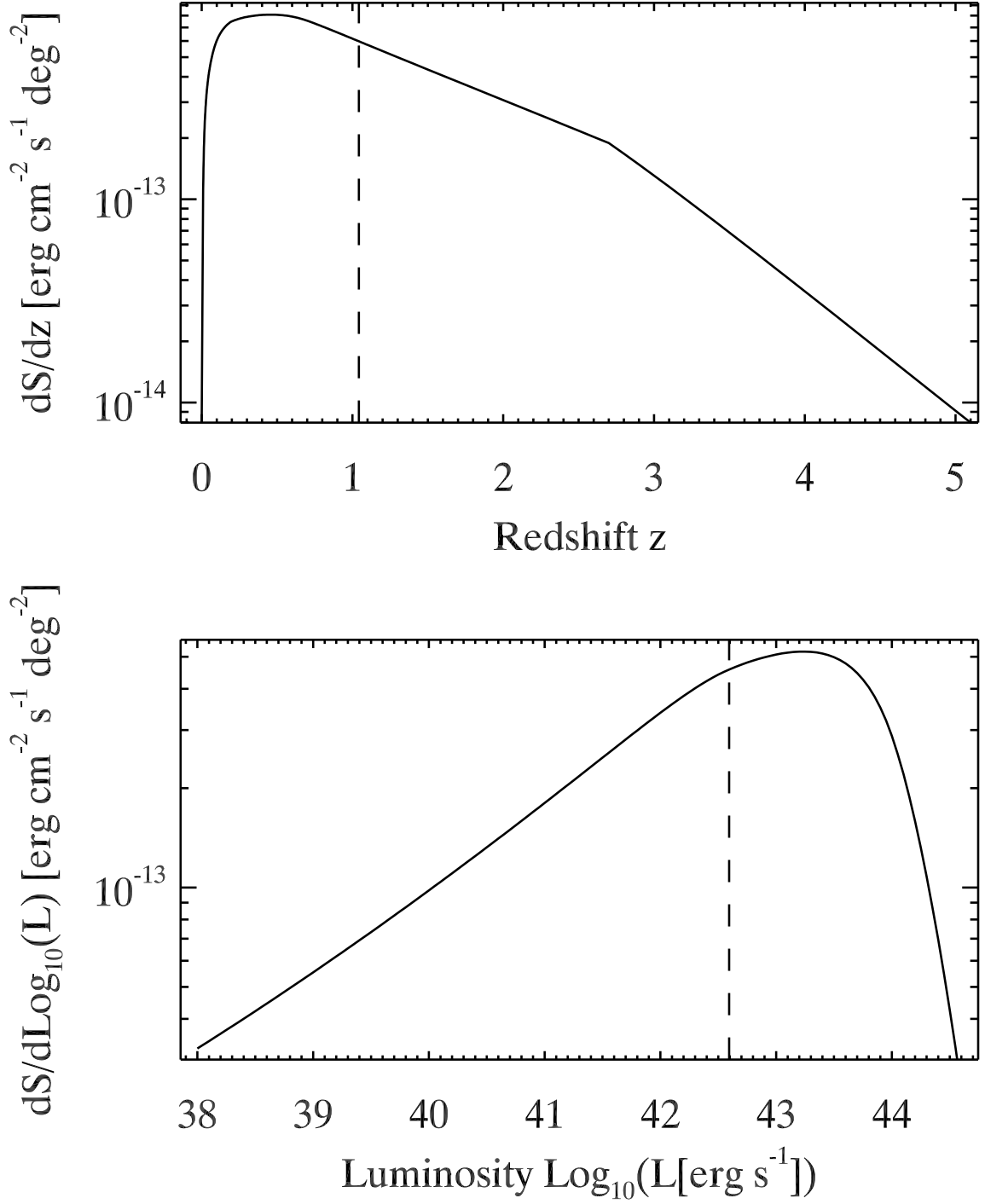


Figure 4.6: Differential flux distribution (alias the flux production rate per solid angle, Eq. 4.6-4.7) for the unresolved AGN in the XBOOTES survey in the 0.5 – 2.0 keV band as a function of redshift (*top*) and luminosity (*bottom*). The dashed lines show the median.

($z > 3$) (also see Sect. 2.5.3 for a discussion of the uncertainties of different XLF of AGN). However, they are nevertheless useful to give us a general idea of main properties of the unresolved AGN population in XBOOTES.

In Fig. 4.6 we show the differential flux distributions of the unresolved AGN in the XBOOTES survey for the 0.5–2.0 keV band. They have a typical redshift (peak) around $z \sim 0.5$ and a median redshift of $z \sim 1.0$. Their median and peak luminosity is around $L \sim 10^{42.6} \text{ erg s}^{-1}$ and $L \sim 10^{43.1} \text{ erg s}^{-1}$, respectively. This means, that the unresolved CXB of XBOOTES enables us to study the clustering signal of relatively low-luminosity AGN around redshift $z \sim 1$, which are otherwise inaccessible with clustering studies of the resolved AGN (e.g. Cappelluti et al. 2012a).

4.4 Brightness fluctuations of the unresolved CXB

In the following we study the brightness fluctuations of the unresolved CXB of XBOOTES by analyzing their power spectrum. In this study we focus on the small angular scales between $\sim 3''$ and $\sim 17'$. In order to achieve the highest possible S/N at these scales, we create the power spectra of all XBOOTES observation and compute their average. This has the advantage in comparison to using a mosaic of all observations for creating the power spectrum that we avoid (for now) further complexities such as overlapping regions, more complex mask, and stronger background flux variations. Further, since we average over more than 100 power spectra, it also enables us to assume a Gaussian error distribution even for the smallest Fourier frequencies, which simplifies error propagation.

The reason for limiting our analysis to a certain interval of angular scales ($\sim 3'' - \sim 17'$) is to reduce the influence of systematics in our analysis. More precise, we ignore all angular frequencies below $10^{-3} \text{ arcsec}^{-1}$ and above 0.3 arcsec^{-1} . The upper limit in angular scales is equal to the side size of the ACIS-I FOV and represents the maximum angular scale, which we are able to measure by using the XBOOTES observations separately. For the frequency regime of $> 0.3 \text{ arcsec}^{-1}$, the source power spectrum is suppressed by more than a factor of 10 due the PSF-smearing (see Fig. 4.7). Since this regime does not contain any more physical information in compare to larger scales, we do not display it for the sake of clarity in Sect. 4.4.4-4.4.7. We nevertheless demonstrate in Sect. 4.4.2-4.4.3 and show in Fig. 4.9 that we are able to describe the power spectrum well also in this frequency regime.

First, we describe our formalism to compute the power spectrum (Sect. 4.4.1). Then we determine the average model of the PSF-smearing for *Chandra*'s ACIS-I (Sect. 4.4.2), which is an essential element for modeling the measured power spectrum. This is followed by several tests of different analysis methods, of the influence

of the instrumental background and for systematics in order to ensure a clear measurement (Sect. 4.4.3-4.4.4). Finally, we present our measurement in Sect. 4.4.5 followed with its interpretation in Sect. 4.4.6-4.4.7.

4.4.1 Formalism

We study the angular brightness fluctuations δF of the unresolved CXB of XBOOTES via Fourier analysis. The Fourier transform of a density field, such as δF , is defined as:

$$\widehat{\delta F}(\mathbf{k}) = \frac{\alpha^2}{(2\pi)^2} \int d^2r \delta F(\mathbf{r}) \exp(-\alpha i \mathbf{r} \cdot \mathbf{k}) . \quad (4.8)$$

For our analysis we use $\alpha = 2\pi$, which connects the angular scale and frequency as $r = k^{-1}$. For the sake of clarity, we insert the value from now on. Due to our observation process the field is transformed from a continuous into a discrete one:

$$\delta F(\mathbf{r}) = \sum_j^N \delta F(\mathbf{r}_j) \delta^D(\mathbf{r} - \mathbf{r}_j), \quad (4.9)$$

where \mathbf{r}_j is the position of the j -th of N image-pixels (where the mask \mathbf{M} equals one), and δ^D is the 2D Dirac delta function. This changes Eq. (4.8) to a 2D discrete Fourier transform:

$$\widehat{\delta F}(\mathbf{k}) = \frac{1}{\sqrt{\Omega}} \sum_j^N \delta F(\mathbf{r}_j) \exp(-2\pi i \mathbf{r}_j \cdot \mathbf{k}) . \quad (4.10)$$

To compute the discrete Fourier transform, we use the FFTW library (v.3.3.3, Frigo & Johnson 2005, <http://www.fftw.org>). The actual *power spectrum* is:

$$\langle |\widehat{\delta F}(k)|^2 \rangle = \frac{2}{n(k)} \sum_j^{n(k)/2} |\widehat{\delta F}(\mathbf{k}_j)|^2 . \quad (4.11)$$

Hereby, the ensemble average $\langle \rangle$ is replaced with the average over all independent Fourier modes $\widehat{\delta F}(\mathbf{k})$ per angular frequency k . There are $n(k)$ Fourier modes within the interval $[k - \Delta k/2, k + \Delta k/2[$ of the 2D Fourier plane, where $\Delta k = L^{-1}$ is defined by the angular size L of the fluctuation map $\delta \mathbf{F}$. This size is defined to be equal for both dimensions ($L_x = L_y$) of the map and to be large enough to embed the entire FOV of an observation. One can analytically approximate the value of $n(k)$

with $2\pi k/\Delta k$ up to the Nyquist-Frequency k_{Ny} . Since the fluctuations map $\delta\mathbf{F}$ is a real quantity, half of the 2D Fourier plan is redundant ($\widehat{\delta F}(\mathbf{k}) = \widehat{\delta F}^*(-\mathbf{k})$, where $*$ indicates the complex conjugate). Therefore, we only have to average in Eq. (4.11) over $n(k)/2$ independent Fourier modes. The number and range of independent Fourier frequencies k is limited by the pixel size Δp , which sets the maximum angular frequency (or minimum angular scale), known as the Nyquist-Frequency, to $k_{\text{Ny}} = (2\Delta p)^{-1}$, and the angular size L of the fluctuation map, which sets the minimum angular frequency (or maximum angular scale) to $k_{\text{min}} = \Delta k = L^{-1}$.

By subtracting the measured power spectrum with our shot-noise estimate P_{Shot} , which is explained and discussed in detail in Sect. 4.4.3, we compute the *shot-noise-subtracted power spectrum*:

$$P(k) = \langle |\widehat{\delta F}(k)|^2 \rangle - P_{\text{Shot}} , \quad (4.12)$$

Note that $P(k)$ is in fact an approximation of the *source power spectrum* $P_{\text{Scr}}(k)$, but we will use both terms as synonyms.

We can assume that each Fourier mode $\widehat{\delta F}(\mathbf{k})$ is approximately Gaussian distributed, since our count-maps \mathbf{C} contain enough pixels with at least one count (typically ~ 600 pixels) and those counts are Poissonian distributed. Based on this we can estimate the statistical uncertainty of $P(k)$ as following:

$$\sigma_P(k) = \sqrt{\frac{2}{n(k)}} \langle |\widehat{\delta F}(k)|^2 \rangle . \quad (4.13)$$

In order to directly compare power spectra of different energy bands we use the *flux-normalized power spectrum*, which we define as:

$$Q(k) = \frac{P(k)}{\left(\Omega^{-1} \sum_j^N F(\mathbf{r}_j) \right)^2} . \quad (4.14)$$

4.4.2 PSF-smearing model

The PSF of *Chandra* will smear out any fluctuation smaller than the size of the PSF, which are caused by source photons. Any fluctuation caused by instrumental background photons are not affected. The described effect is called *PSF-smearing* (also known as beam smearing) and leads to a drop of the source power spectrum $P_{\text{Scr}}(k)$ amplitude at small scales (e.g. Fig. 4.14). Since the size and shape of the PSF changes with CCD offset angle (θ) and position (ϕ) (see Fig. 4.1), one can not derive a simple analytical expression to describe this effect. Therefore, we use an

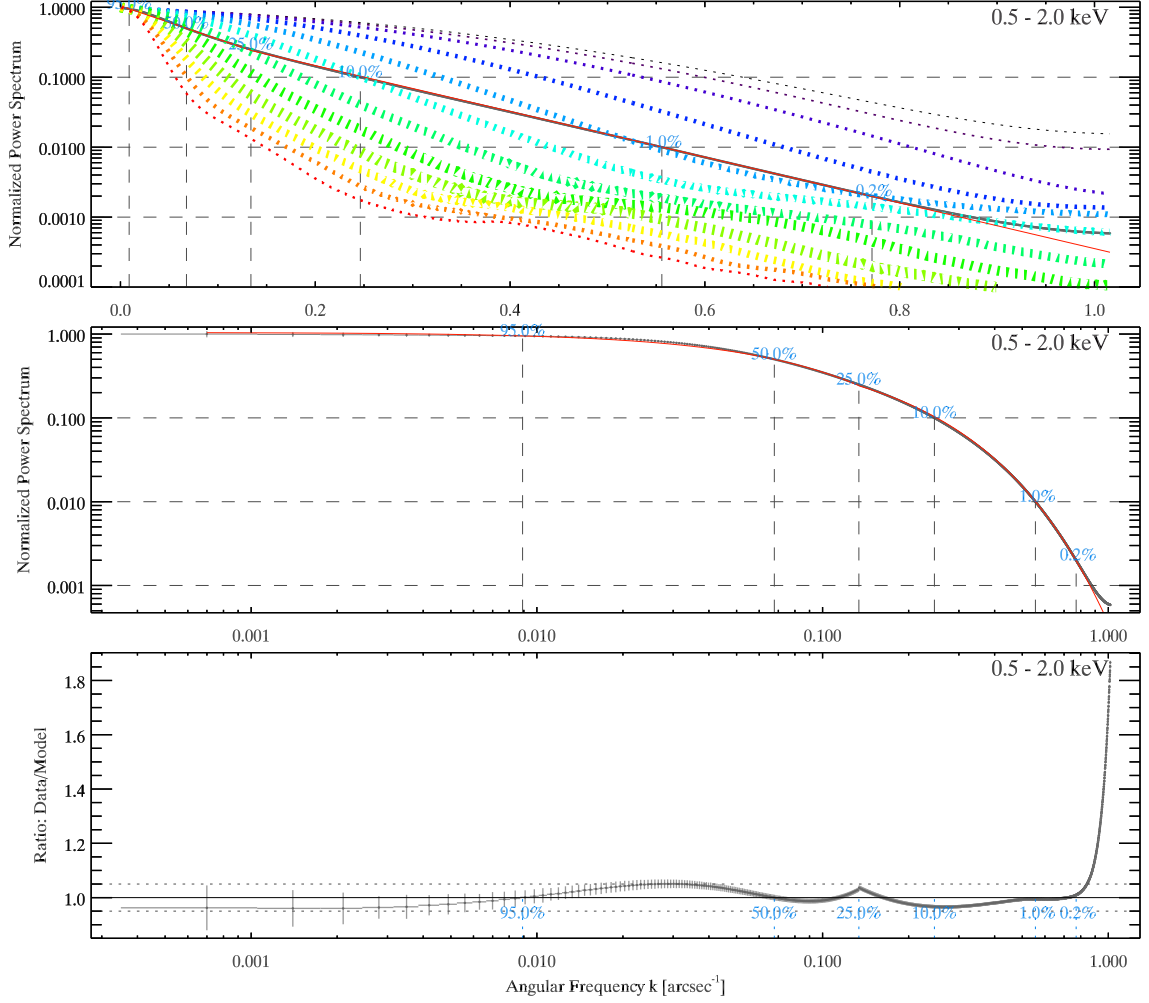


Figure 4.7: *Two upper panels:* Average power spectrum of the PSF of *Chandra* ACIS-I (gray crosses, average of all offset angles and CCDs) and our empirical PSF-smearing model (red curve, Eq. 4.15). In the first panel we also show the average PSF power spectrum (dotted curves) for different offset angles from 0' (black) to 11' (red), where the thickness of the curve indicates the weight for the weighted average over all offset angles. *Lower panel:* Ratio between the measurement and our PSF-smearing model.

Table 4.8: Parameters of our PSF model (Eq. 4.15)

Symbol	Value	Unit
k_b	0.134	arcsec^{-1}
α_1	-4.740	arcsec
β_1	0.020	-
α_2	-3.270	arcsec
β_2	-0.180	-

empirical approach and compute an average PSF power spectrum from the measured PSF power spectra of our PSF simulations in Sect. 4.2.2 for the 0.5 – 2.0 keV band. Thereby, we first average over all CCDs and then compute a weighted average over all offset angles. As weight serves the surface area of the annulus of each offset angle. We show the average PSF power spectrum in Fig. 4.7 (grey crosses) along with our simple empirical *PSF-smearing model* (red curve), which is a broken exponential function in the base of ten:

$$\log_{10} [P_{\text{PSF}}(k)] = \begin{cases} \alpha_1 k + \beta_1 & \text{for } k \leq k_b, \\ \alpha_2 k + \beta_2 & \text{for } k \geq k_b. \end{cases} \quad (4.15)$$

The parameters of the model are shown in Table 4.8. Our model is able to describe the PSF power spectrum up to a frequency of $k \approx 0.9 \text{ arcsec}^{-1}$ to an accuracy of $\approx 5\%$, which is sufficient for the given S/N of our fluctuation measurement. We can also see in Fig. 4.9 that our PSF-smearing model is able to describe well the measured power spectrum down very small angular scales.

4.4.3 Photon shot noise

When measuring the angular fluctuations of the unresolved CXB via Fourier analysis, we have to take into account that the source power spectrum $P_{\text{Scr}}(k)$ is superimposed by the *photon shot noise* P_{Shot} (Eq. 4.12) in our measured power spectrum $\langle |\widehat{\delta F}(k)|^2 \rangle$. In the ideal case the shot noise is a flat component and independent of any Fourier frequency. Its amplitude is proportional to the total number of counts (N_{Counts}) and it limits the maximum theoretical S/N of $P_{\text{Scr}}(k)$, in particular for the low count regime: $\text{S/N} \propto (1 - N_{\text{Counts}}^{-1})$.

4.4.3.1 Estimators

We present here three different photon shot-noise estimators.

Analytical estimate When we are ignoring the effects of the mask (Sect. 4.4.4.1) and the vignetting (Sect. 4.4.4.4) in the measured power spectrum, we can use a simple analytical expression to estimate the shot-noise amplitude, which we derive here. Based on our definition in Sect. 4.4.1, we can write the measured power spectrum as following:

$$\begin{aligned} \left\langle |\widehat{\delta F}(\mathbf{k})|^2 \right\rangle &= \frac{1}{\Omega} \left\langle \left(\sum_j^N \delta F_j e^{-2\pi i \mathbf{r}_j \mathbf{k}} \right) \left(\sum_l^N \delta F_l e^{+2\pi i \mathbf{r}_l \mathbf{k}} \right) \right\rangle \\ &= \frac{1}{\Omega} \left\langle \sum_j^N \delta F_j^2 \right\rangle + \frac{1}{\Omega} \left\langle \sum_j^N \sum_{l \neq j}^{N-1} \delta F_j \delta F_l e^{-2\pi i \mathbf{k}(\mathbf{r}_j - \mathbf{r}_l)} \right\rangle. \end{aligned}$$

The second term represent the actually source power spectrum $P_{\text{Scr}}(k)$, whereas the first term is the shot noise. We can reduce the first term further to:

$$\begin{aligned} \left\langle \sum_j^N \delta F_j^2 \right\rangle &= \sum_j^N \langle \delta F_j^2 \rangle = \sum_j^N (\langle F_j^2 \rangle - \langle F_j \rangle^2) \\ &= \sum_j^N \sigma^2(F_j) = \sum_j^N \frac{\sigma^2(C_j)}{E_j^2} \approx \sum_j^N \frac{C_j}{E_j^2}. \end{aligned}$$

Here, we use the fact that $\delta F_j = F_j - \langle F_j \rangle$ and $F_j = C_j/E_j$ and that the variance $\sigma^2(x)$ of a Poissonian-distributed quantity x is equal to its mean. Since we do not know the actually mean of C_j we approximate it with its own value. Based on this derivation, we define our shot noise estimate as following:

$$P_{\text{Shot}} = \frac{1}{\Omega} \sum_j^N \frac{C(\mathbf{r}_j)}{E^2(\mathbf{r}_j)}. \quad (4.16)$$

When we are using the average exposure map (Eq. 4.1) for computing the fluctuation map $\delta \mathbf{F}$, the definition changes to:

$$P_{\text{Shot}} = \frac{1}{\Omega} \frac{\sum_j^N C(\mathbf{r}_j)}{\langle E \rangle^2}.$$

Hereby is $\langle E \rangle = \sum_j^N E(\mathbf{r}_j)/N$. We refer to both definitions as the *analytical shot-noise estimate*, which is our default shot-noise estimate.

Note that this definition is only valid for the total-count map $\mathbf{C}_X^{\text{Obs.}}$ (Sect. 4.2.4). For the background-subtracted map $\mathbf{C}_X^{\text{Scr.}}$ (Eq. 4.3) one also has to take into account the additional shot noise by the stowed-background map $\mathbf{C}^{\text{sto.BKG}}$:

$$P_{\text{Shot}}^{(\text{Scr.})} = P_{\text{Shot}} + P_{\text{Shot}}^{(\text{Obs.BKG})} ,$$

with (using Eq. 4.2)

$$P_{\text{Shot}}^{(\text{Obs.BKG})} = \frac{\sum_j^N C_j^{\text{sto.BKG}}}{\Omega \langle E \rangle^2} \cdot \left(\frac{\sum_j^N C_{9.5-12.0 \text{ keV}}^{\text{Obs.}}}{\sum_j^N C_{9.5-12.0 \text{ keV}}^{\text{sto.BKG}}} \right)^2 . \quad (4.17)$$

High-frequency based estimate The *high-frequency based shot-noise estimate* uses the fact that at very high frequencies the measured power spectrum converges eventually to the shot-noise amplitude due to the PSF-smearing (Sect. 4.4.2). Hence, we can estimate the shot noise by taking the average ($P_{\text{Shot}}^{(\text{HF})}$) of the power spectrum for a frequency interval $[k_{\text{min}}^{(\text{HF})}, k_{\text{max}}^{(\text{HF})}]$, where the upper limit is the Nyquist-Frequency: $k_{\text{max}}^{(\text{HF})} = k_{\text{Ny}}$. The lower limit of this interval is somewhat arbitrary. We use $k_{\text{min}}^{(\text{HF})} = k_{\text{Ny}} \times 0.80 \approx 0.81 \text{ arcsec}^{-1}$. In this case, the interval encapsulates angular scales of the size ≈ 2.5 ACIS-I CCD pixels or smaller ($\leq 1.2''$). At these scales the amplitude of the source power spectrum is suppressed by more than 500 times due to PSF-smearing (Fig. 4.7) and the shot-noise-subtracted power spectrum in the 0.5–2.0 keV band is more than ~ 2000 times smaller than the shot-noise amplitude itself (using the analytical shot-noise estimate). Therefore, we assume (and confirm in Fig. 4.9) that within this frequency interval we can neglect the source power spectrum. We also tested larger and smaller values for $k_{\text{min}}^{(\text{HF})}$ and find that $k_{\text{min}}^{(\text{HF})} = k_{\text{Ny}} \times 0.80$ is a good compromise between being small enough to neglect the source power spectrum and being large enough to ensure good statistic in computing an average value. It is further reassuring that our results of the fluctuation analysis in Sect. 4.4.5–4.4.7 do not depend strongly on the choice of $k_{\text{min}}^{(\text{HF})}$. We could have in principle also used $k_{\text{min}}^{(\text{HF})} = k_{\text{Ny}} \times 0.40 \approx 0.41 \text{ arcsec}^{-1}$ and we would not see a significant difference of the shot-noise-subtracted power spectrum for frequencies of $\lesssim 0.3 \text{ arcsec}^{-1}$ in compare to using the analytical shot-noise estimate.

However, for *Chandra* ACIS-I our definition of the high-frequency based estimate works only optimal for an image-pixel-binning of one. Already for an image-pixel-binning of two ($\Delta p = 2 \times 0.492''$) the corresponding Nyquist-Frequency ($\approx 0.5 \text{ arcsec}^{-1}$) would be smaller than our $k_{\text{min}}^{(\text{HF})}$. In the case, we would have to use

a different frequency interval, where the source power spectrum is not neglectable anymore. This would lead ultimately to an overestimation of the shot noise and a removal of a significant fraction of source power spectrum for the shot-noise-subtracted power spectrum. The effect has the worst impact at the high-frequency part of the measured power spectrum and might even be neglectable at very small frequencies given a sufficiently high amplitude of the source power spectrum in this frequency regime.

For the shot-noise-subtracted power spectrum, we take the statistical uncertainties of the measured power spectrum and of $P_{\text{Shot}}^{(\text{HF})}$ into account and use Gaussian error propagation.

As we show in the Sect. 4.4.3.2, the high-frequency based estimate works as good as the analytical estimate and would lead in principle also to the same scientific results. The general concern with this estimate is that it works only optimal for an image-pixel-binning of one and that it has the potential to overestimate the shot-noise if there is still a significant source power spectrum in the frequency interval $[k_{\text{min}}^{(\text{HF})}, k_{\text{max}}^{(\text{HF})}]$. Nevertheless, the estimate serves as an good independent test to verify that the analytical estimate is a good shot-noise estimate.

Observation-splitting based estimate The *observation-splitting shot-noise based estimate*, also called the $A - B$ technique, was introduced by Kashlinsky et al. (2005) for infrared data and was also used for X-ray data (Cappelluti et al. 2012b, 2013; Helgason et al. 2014). Hereby, one splits the observation into even and odd time-frames (or events) and creates a fluctuations map for each subset ($\delta\mathbf{F}_A$ for even frames and $\delta\mathbf{F}_B$ for odd frames). The difference between those two maps should remove in principle any source power spectrum or stable instrumental effects, because both subsets were observed almost simultaneously, and therefore only contain the random noise of the observation. Hence, the measured power spectrum $P_{\text{Shot}}^{(\text{OS})}$ of the difference $\delta\mathbf{F}_D = (\delta\mathbf{F}_A - \delta\mathbf{F}_B)/2$ should represent the shot noise for this observation. For the shot-noise-subtracted power spectrum, we take the statistical uncertainties of the measured power spectrum and of $P_{\text{Shot}}^{(\text{OS})}$ into account and use Gaussian error propagation.

As we show in the Sect. 4.4.3.2, observation-splitting based estimate gives in average consistent results in comparison to the analytical estimate. However, a major concern with this estimate is that this estimate is itself subject to noise. Combined with correlations between different Fourier modes caused by the mask effect this leads to appearances of irregularities in the resulting shot-noise-subtracted power spectrum. This can be seen for instance in Fig. 4.9 (lower panels, green crosses) for in the interval of $0.5 \lesssim k[\text{arcsec}^{-1}] \lesssim 0.7$, where the observation-splitting

based estimate leads to a significant underestimation of the power spectrum. If it is not the aim to measure the shape of the power spectrum to high accuracy, one might ignore such a behavior. However, it also makes the observation-splitting based estimate somewhat less reliable in comparison to the analytical or high-frequency based estimate. Therefore, we did not use it as our default estimate but it serves as an independent test to verify that the analytical and high-frequency based estimate are good shot-noise estimates.

4.4.3.2 Evaluation

In the following, we compare our different shot-noise estimates. Thereby, we focus mostly on the high frequency part of the power spectrum, where the source power spectrum is of the order of the shot-noise amplitude or smaller. For a better visualization, we show the power spectrum only in linear scales and use a linear binning, although all calculations were performed with the unbinned power spectrum.

The 9.5–12.0 keV band The power spectrum in the 9.5–12.0 keV band is ideal for evaluating our shot-noise estimates. Since *Chandra* has no effective area in this energy band, we are only able to detect instrumental background counts but no source counts, and since the instrumental background counts are Poissonian-distributed, we know that their power spectrum is equal to their shot noise itself. If we compare this expectation with our shot-noise-subtracted power spectrum in Fig. 4.8, we can see that indeed also these power spectra appear in average flat.

We can also see that the different shot-noise-subtracted power spectra fluctuate strongly around zero. The behavior is caused by the mask effect leading to correlations of different Fourier modes, which is described more in detail in Sect. 4.4.4.1. To also have an independent method of estimating the statistical uncertainty for a frequency bin, we also show for the analytical estimate the standard deviation of the sample mean for each bin (black error-bars) in Figs. 4.8-4.9. There are in general only slightly higher than those from the statistical uncertainty of the power spectrum (Eq. 4.13).

The very first frequency bin (linear binning) contains also a real signal at very large scales, which also can be seen for the ACIS-I stowed-background map (see Sect. 4.4.4.3 and Fig. 4.12). Therefore, we exclude the frequencies of $< 0.02 \text{ arcsec}^{-1}$ in the following evaluation.

To statistically evaluate the different estimates we compute the sample mean and RMS deviation of the quantity: $Z(k) = P(k)/\sigma_{P(k)}$. Since the unbinned power spectrum has more than 1000 Fourier frequencies, the average of $Z(k)$ is Gaussian

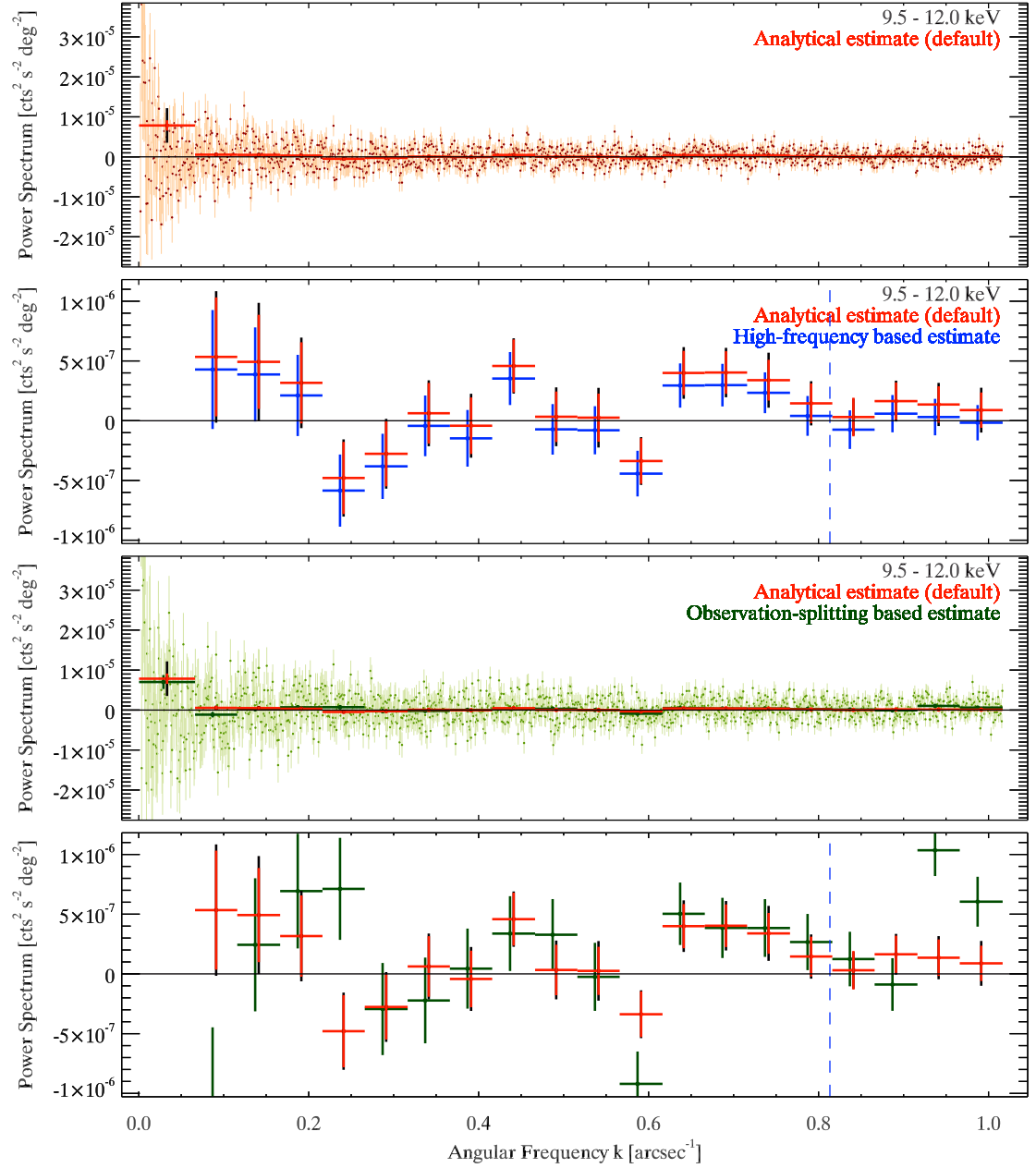


Figure 4.8: Shot-noise-subtracted power spectrum $P(k)$ (Eq. 4.12) in the 9.5 – 12.0 keV band. Red crosses (all panels): using the analytical shot-noise estimate (default) P_{Shot} (Eq. 4.16). Blue crosses (upper panels): using the high-frequency based shot-noise estimate $P_{\text{Shot}}^{(\text{HF})}$ (Sect. 4.4.3.1). Green crosses (lower panels): using the observation-splitting based estimate $P_{\text{Shot}}^{(\text{OS})}$ (Sect. 4.4.3.1). Colored error-bars: Standard deviation due to error propagation of all binned frequencies. Black error-bars: Standard deviation of the sample mean of all binned frequencies. Vertical blue dashed line: lower limit $k_{\text{min}}^{(\text{HF})}$ to estimate $P_{\text{Shot}}^{(\text{HF})}$.

distributed and is in the ideal case zero with a RMS deviation of one. We measure a RMS deviation of ≈ 1.14 for all shot-noise estimates. The value is reasonable close to the expectation. We measure a sample mean of 0.05 ± 0.03 , 0.00 ± 0.03 and 0.08 ± 0.03 for the analytical, high-frequency based, and observation-splitting based estimate, respectively. This means that the first two shot-noise-subtracted power spectra are consistent with zero within two standard deviations. Therefore, they are in average good estimators. We choose the analytical estimate as our default estimator, because it uses the smallest amount of assumptions. Additional difficulties of the alternative estimators are discussed in their description of Sect. 4.4.3.1.

The 0.5 – 2.0 keV band To demonstrate that our shot-noise estimates also work properly for the brightness fluctuations of the unresolved CXB of XBOOTES, we show the shot-noise-subtracted power spectrum $P(k)$ (Eq. 4.12) also for the 0.5 – 2.0 keV band for all our shot-noise estimates in Fig. 4.9 (red, blue and green crosses for the analytical, high-frequency based, and observation-splitting based estimate, respectively). They show deviations, which appear to be statistically significant. As discussed above, these deviations are caused by the mask effect leading to correlations of nearby Fourier modes, in particular for the observation-splitting based. We also show our AGN clustering model (purple curve, Sect. 4.4.6), which is multiplied with our PSF-smearing model (Sect. 4.4.2). We can see that for all estimates the shot-noise-subtracted power spectrum fluctuates around zero for $\gtrsim 0.8 \text{ arcsec}^{-1}$ and follow our power spectrum model (purple curve) for smaller frequencies. This also verifies that we can indeed neglect the source power spectrum for the frequencies of $> k_{\text{Ny}} \times 0.80 \approx 0.81 \text{ arcsec}^{-1}$, which is used as an assumption to compute the high-frequency based estimate (Sect. 4.4.3.1) for the frequency interval of $[k_{\text{Ny}} \times 0.80, k_{\text{Ny}}]$. Fig. 4.9 further demonstrates that our PSF-smearing model is able to describe the measured power spectrum up to the highest frequencies, where the source power spectrum becomes eventually neglectable.

We should note that an incorrectly estimated shot-noise amplitude has its largest impact at the highest frequencies of the power spectrum. This leads to the fact that for frequencies of $\lesssim 0.4 \text{ arcsec}^{-1}$ all shot-noise estimates do not show a very significant difference in the shot-noise-subtracted power spectrum. This is reassuring, because it illustrates that our results (Sect. 4.4.5-4.4.7) are robust against the exact choice of our shot-noise estimate.

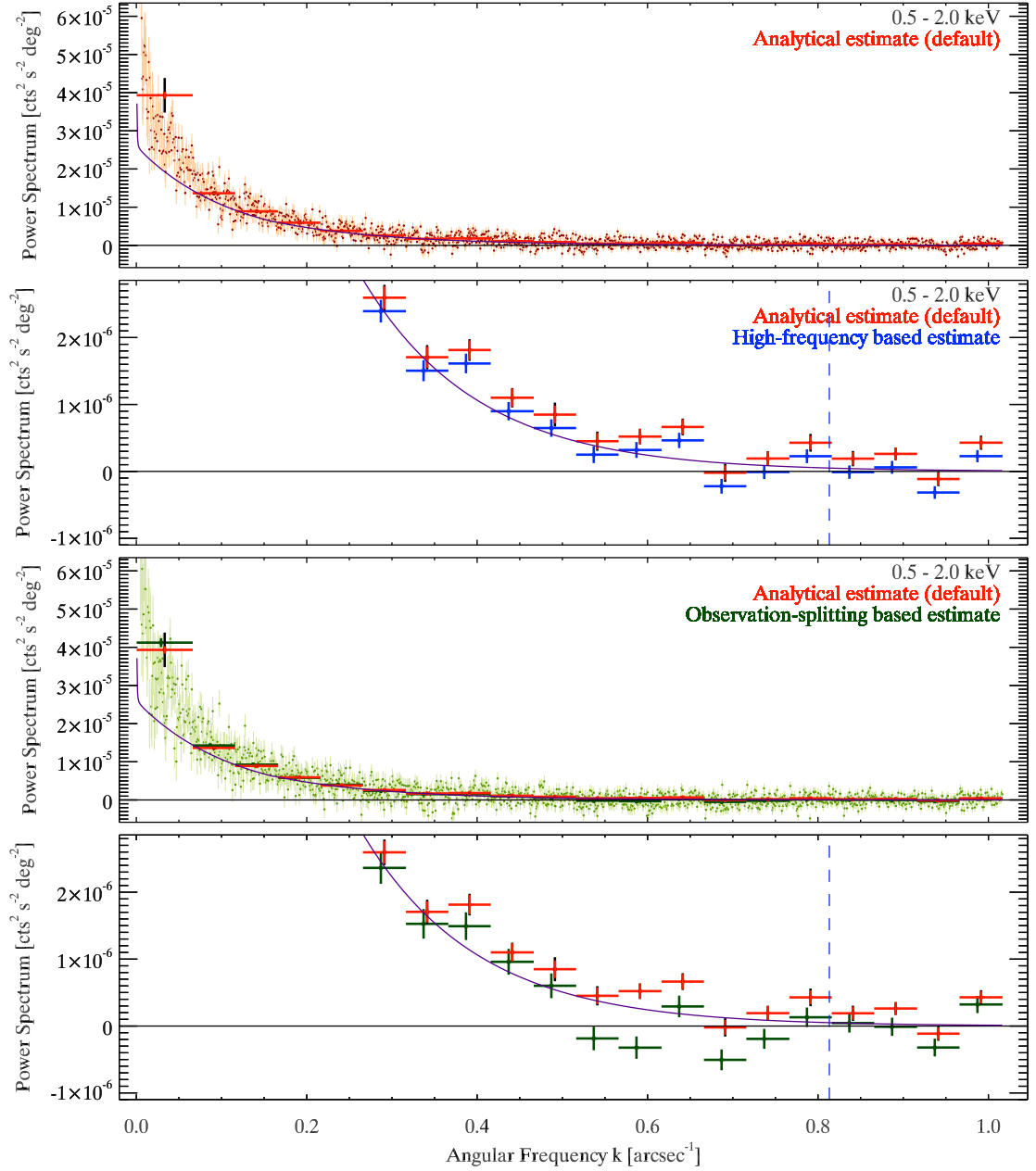


Figure 4.9: Same as Fig. 4.8 for the 0.5 – 2.0 keV band. Purple solid curve: AGN clustering model (Sect. 4.4.6), which is multiplied with our PSF-smearing model (Sect. 4.4.2). We see that our PSF-smearing model can describe the measured power spectrum up to high frequencies.

4.4.4 Tests for systematics

Here we present several tests for systematics in order to show that our results (Sect. 4.4.5-4.4.7) are robust against them.

4.4.4.1 Mask effect

Our measured power spectrum (e.g. Fig. 4.14) is actually a convolution of the true power spectrum with a window function, because we are using a masked image, which has a finite size defined by the ACIS-I FOV and also masked out region due removed resolved sources (Sect. 4.2.2) and CCD gaps and bad pixels. This convolution alters the power spectrum and leads to correlation of adjacent Fourier frequencies. We summarized these modifications with the term *mask effect*. Here, we estimate how the mask effect distorts the source power spectrum in the considered frequency interval of $10^{-3} - 0.3 \text{ arcsec}^{-1}$.

We do this estimation by using a model of our measured power spectrum (0.5 – 2.0 keV, Fig. 4.14). Since we do not have a complete model to describe our measured power spectrum (Sect. 4.4.7), we use an ad-hoc model for this purpose, which consists of our AGN clustering model (Sect. 4.4.5) and a simple powerlaw to describe the excess, which all are multiplied with our PSF-smearing model (Sect. 4.4.2). The convolution is done by creating a two-dimensional Fourier-transform from the one-dimensional power spectrum model with random, uniformly distributed phases. From this we make an inverse Fourier-transform to compute a real image, which is then multiply with the mask of each observation. We compute the power spectrum for each of these images as described in Sect. 4.4.1 and compute the average from all these power spectra. We perform this simulation 50 times, which is sufficient for a good convergence, and compute the average of these simulations.

We show the result in Fig. 4.10. Here, we see that due the mask effect the power spectrum is suppressed by only 40 % at the lowest frequency bin and frequencies of $\gtrsim 8 \times 10^{-3} \text{ arcsec}^{-1}$ are not affected at all. This means, if our ad-hoc model would be indeed a good representation of our measured power spectrum, that the true power spectrum would have a slightly larger amplitude at low frequency than what we currently measure. Of course, this effect should be accurately taken into account when we would fit the measured power spectrum with an appropriate theoretical model. However, it reassuring to notice that the overall shape of measured power spectrum does not change very significantly.

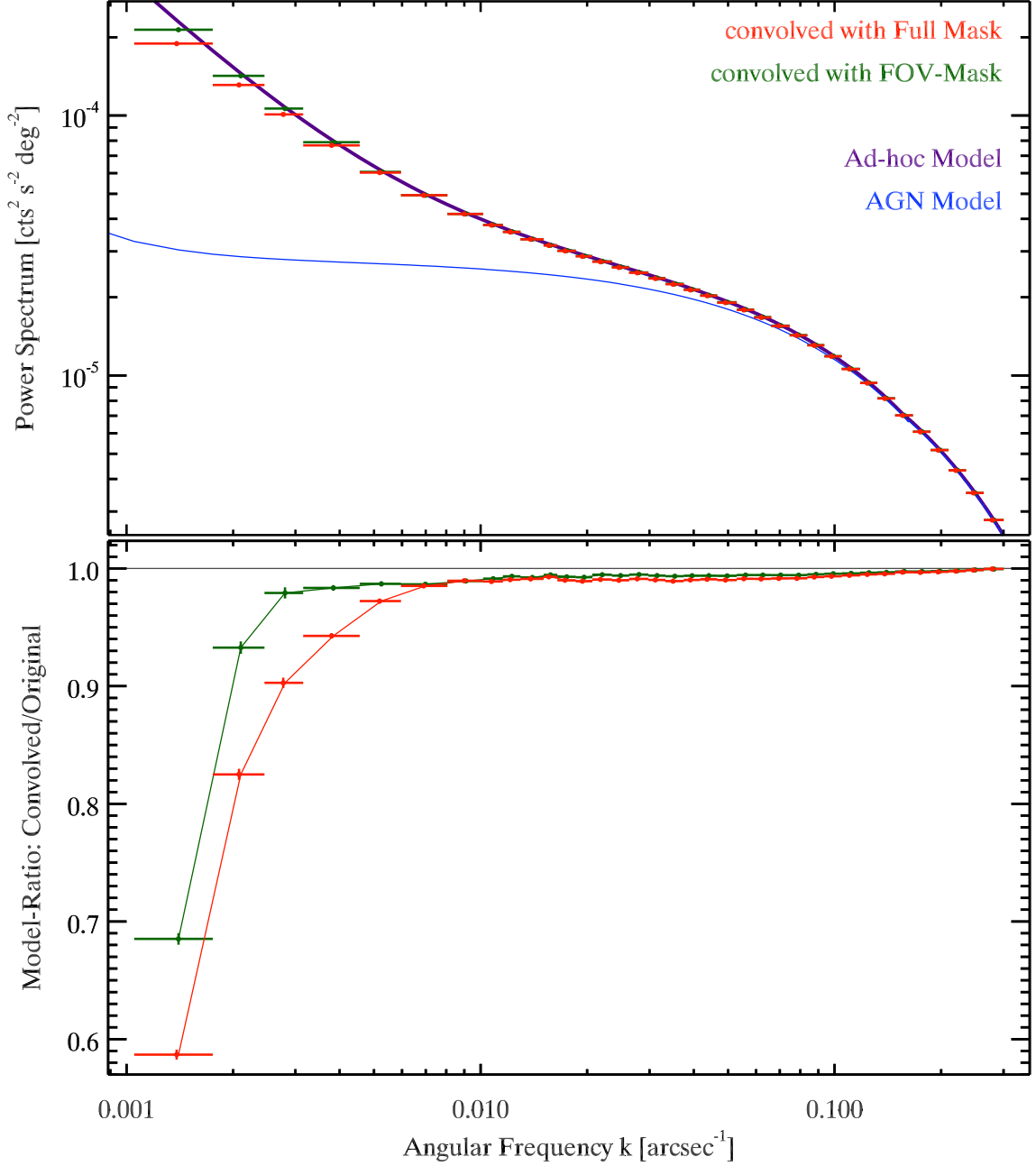


Figure 4.10: To demonstrate the mask effect we use an ad-hoc model (purple curve), which represents our measured power spectrum (0.5 – 2.0 keV, Fig. 4.14), and convolve it with the full mask (red crosses, regions of resolved sources masked out) and only with the ACIS-I FOV-Mask (green crosses, regions of resolved sources *not* masked out) of all XBOOTES observations. Also shown our AGN clustering model (blue curve, Sect. 4.4.6).

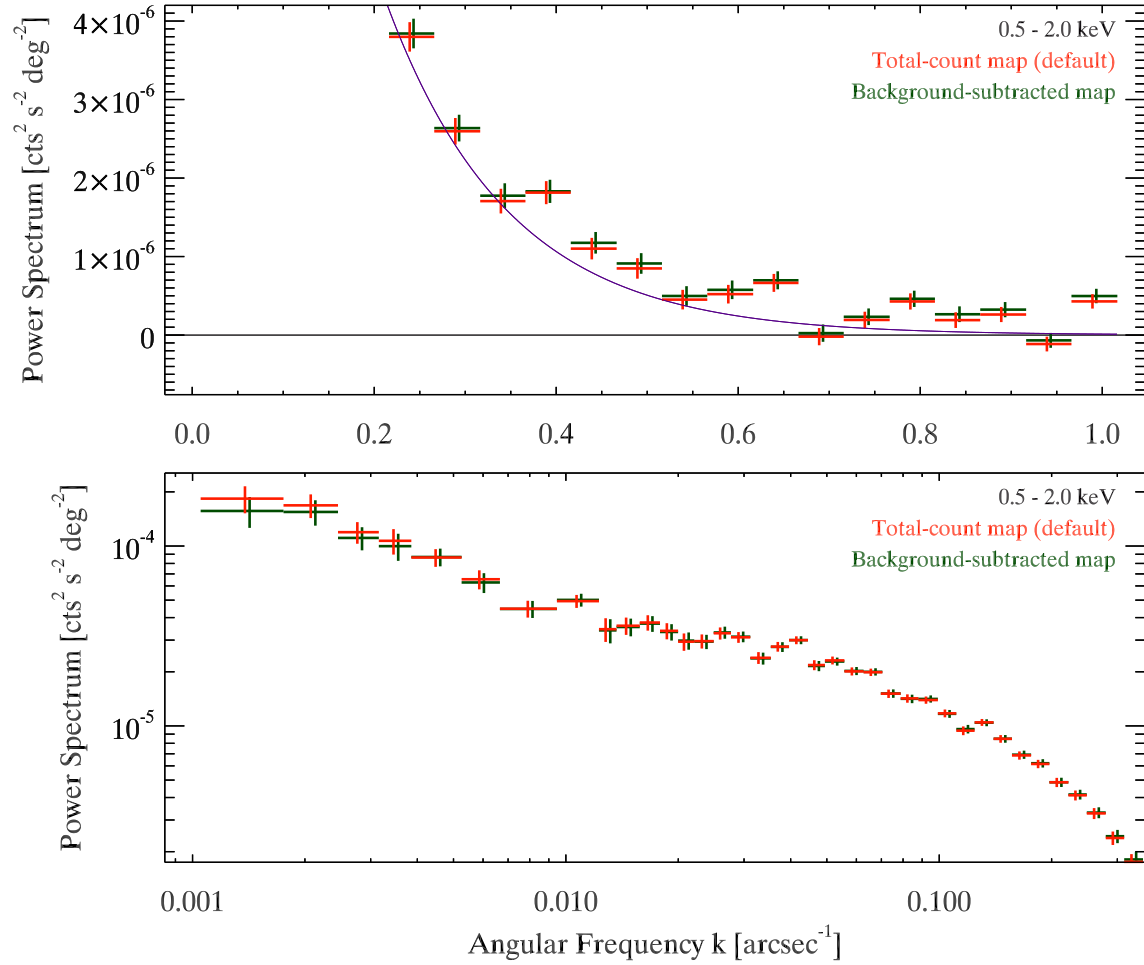


Figure 4.11: Shot-noise-subtracted power spectrum $P(k)$ (Eq. 4.12) in the 0.5 – 2.0 keV band computed with the total-count map $\mathbf{C}_X^{\text{Obs.}}$ (red crosses) and the background-subtracted map $\mathbf{C}_X^{\text{Scr.}}$ (green crosses, Eq. 4.3, Sect. 4.2.4). Purple solid curve: AGN clustering model (Sect. 4.4.6) multiplied with our PSF-smearing model (Sect. 4.4.2). We can see that for the shown frequencies range, there is no significant difference in computing the power spectrum with those two maps.

4.4.4.2 Total-count map vs. background-subtracted map

For our analysis we are using the total-count map $\mathbf{C}_X^{\text{Obs.}}$ (Sect. 4.2.4) instead of the background-subtracted map $\mathbf{C}_X^{\text{Scr.}}$ (Eq. 4.3) to create the fluctuation map $\delta\mathbf{F}_X$ (Eq. 4.4). In Fig. 4.11 we show that both power spectra in the $0.5 - 2.0$ keV band do not show any significant differences for the considered angular scales of $\sim 3'' - \sim 17'$ ($10^{-3} - 0.3 \text{ arcsec}^{-1}$). This shows that it is not important, which type of map we use for our analysis and it also means that adding the fluctuations of the stowed background to the total-count map (Eq. 4.16) does not alter the power spectrum significantly for these angular scales.

4.4.4.3 Instrumental, and stowed-background map

In Fig. 4.12 we compare the flux-normalized power spectra $Q(k)$ (Eq. 4.14) of the total-count map $\mathbf{C}_X^{\text{Obs.}}$ (red crosses, Sect. 4.2.4) and the stowed-background map $\mathbf{C}^{\text{sto.BKG}}$ (blue crosses, Sect. 4.2.4) in the $0.5 - 2.0$ keV and $1.0 - 2.0$ keV band and of the total-count map in the $9.5 - 12.0$ keV band (green crosses). We can see that power spectra of the stowed-background map and of the $9.5 - 12.0$ keV band are very similar, which confirms the assumption that the fluctuations in these two maps have the same instrumental origin. They both show a significant clustering signal for scales of $\gtrsim 50''$ ($\lesssim 0.02 \text{ arcsec}^{-1}$), which is probably caused from the surrounding inhomogeneous telescope structure of the *Chandra* ACIS-I CDDs. Nevertheless, we can conclude that the power spectrum of the instrumental background is still more than an order of magnitude smaller than the source power spectrum (red crosses).

4.4.4.4 Vignetting

Instrumental background counts are not affected by vignetting but if we create the fluctuation map in the $9.5 - 12.0$ keV band with the original exposure map instead of the average exposure map (Sect. 4.2.1), we are able to simulate the effect of vignetting onto a flat shaped power spectrum. When comparing the two power spectra in the $9.5 - 12.0$ keV band (not shown), where one uses the original and average exposure map, we can not see any significant difference on all scales. Source counts are indeed affected by vignetting, but if we compare the power spectra in the $0.5 - 2.0$ keV band in Fig. 4.13, where we use the original (red crosses) and average exposure map (blue crosses), we neither can see any significant differences on all scales. Hence, we can conclude that the distortions of the power spectrum due to vignetting can be neglected for our analysis at the given S/N.

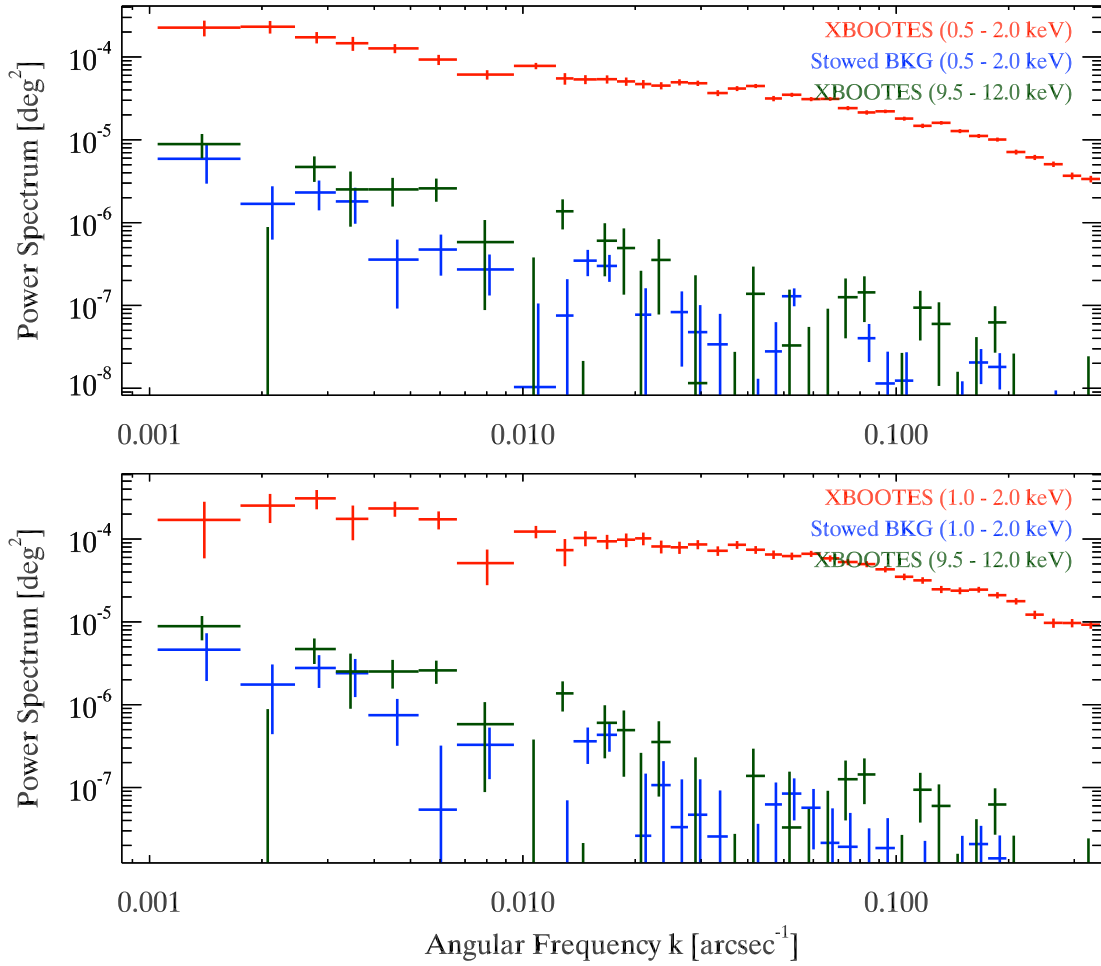


Figure 4.12: Flux-normalized, shot-noise-subtracted power spectra $Q(k)$ (Eq. 4.14) of the total-count map $\mathbf{C}^{\text{Obs.}}$ (red crosses) and of the ACIS-I stowed-background map $\mathbf{C}^{\text{sto.BKG}}$ (blue crosses, Sect. 4.2.4) for the indicated energy bands, and for the total-count map $\mathbf{C}^{\text{Obs.}}$ for the 9.5 – 12.0 keV band (green crosses).

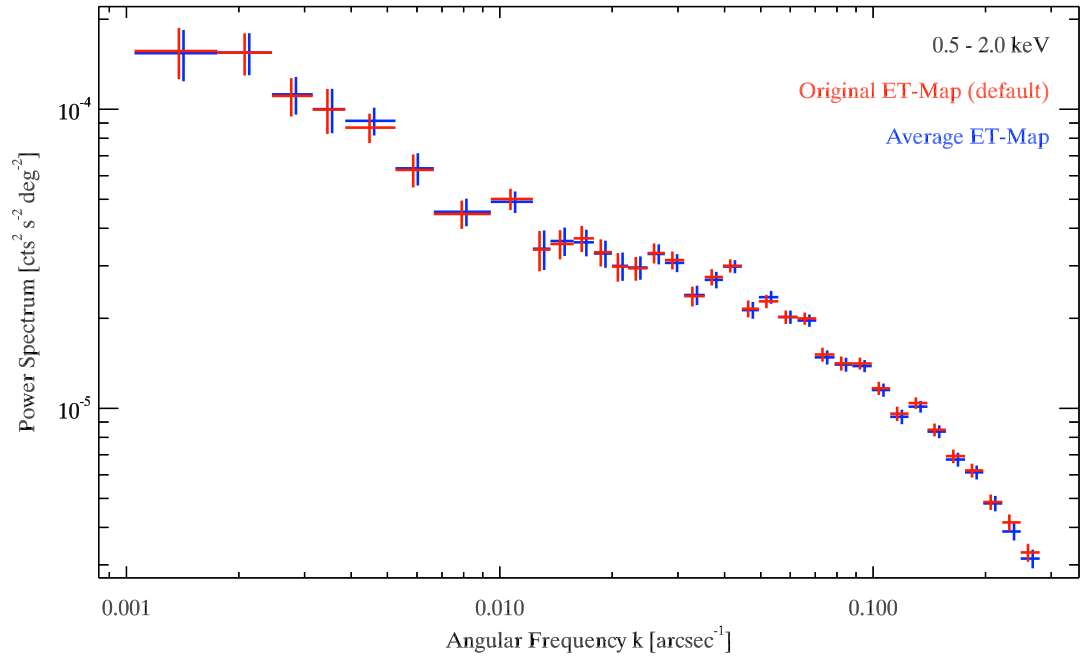


Figure 4.13: Shot-noise-subtracted power spectra $P(k)$ (Eq. 4.12) of the total-count map in the case of using the original (red crosses) and average (blue crosses) exposure map (Sect. 4.2.1).

4.4.5 Power spectrum of the brightness fluctuations

In Fig. 4.14 we present the power spectrum of the brightness fluctuations of the unresolved CXB of XBOOTES in the $0.5 - 2.0$ keV band. In the following, we focus our analysis of these fluctuations in the $1.0 - 2.0$ keV band. The corresponding power spectrum is shown in Fig. 4.15. The reason for this is to minimize the contamination of fluctuations caused by unresolved emission of our own Galaxy, since Galactic emission can be neglected for > 1 keV based on our spectral fit of the unresolved CXB (Sect. 4.3). There are indications by previous studies (Śliwa et al. 2001; Cappelluti et al. 2013; Helgason et al. 2014) that Galactic emission may contribute significantly to the brightness fluctuations, in particular for larger scales ($\gtrsim 2'$). We also detect a significant higher amplitude of the power spectrum for the $0.5 - 1.0$ keV band in comparison to our AGN clustering model (Sect. 4.4.7.3). If these additional fluctuations have indeed Galactic origin, than fluctuation analysis with XBOOTES would offer us a new observational constraint on Galactic emission. In order to give this topic its full attention, we will analyze the fluctuations of the unresolved CXB for < 1.0 keV in our subsequent study, where we should be able to compare our results with the measurement of Śliwa et al. (2001) since we will also measure larger scales, but ignore it in the current study.

We show in our spectral analysis in Sect. 4.3 that the energy spectrum of the unresolved CXB is consistent with being dominated by AGN. Further, previous studies (Cappelluti et al. 2013; Helgason et al. 2014) with a deeper survey (more than 30 times in sensitivity) show that the clustering signal of AGN still dominates the power spectrum of the brightness fluctuations of the unresolved CXB for angular scales of $\lesssim 2'$. Therefore, we focus first on the clustering signal of AGN and compare it with our measurement (Sect. 4.4.6) before investigating additional clustering sources in order to explain our measured power spectrum (Sect. 4.4.7).

We note that in the following all clustering models are multiplied by our PSF-smearing model (Sect. 4.4.2).

4.4.6 Clustering signal of unresolved AGN

The clustering signal of AGN can be described with the halo model approach (e.g. Cooray & Sheth 2002), which results in a power spectrum model of two terms: The one-halo-term (1H) describes the correlation of AGN, which reside in the same dark matter halo (DMH), whereas the two-halo-term (2H) describes the correlation of AGN from different DMHs. We build our AGN clustering model based on the assumption, supported by current clustering measurements of resolved AGN, that there is only one AGN per DMH (e.g. Miyaji et al. 2011; Allevato et al. 2012) and

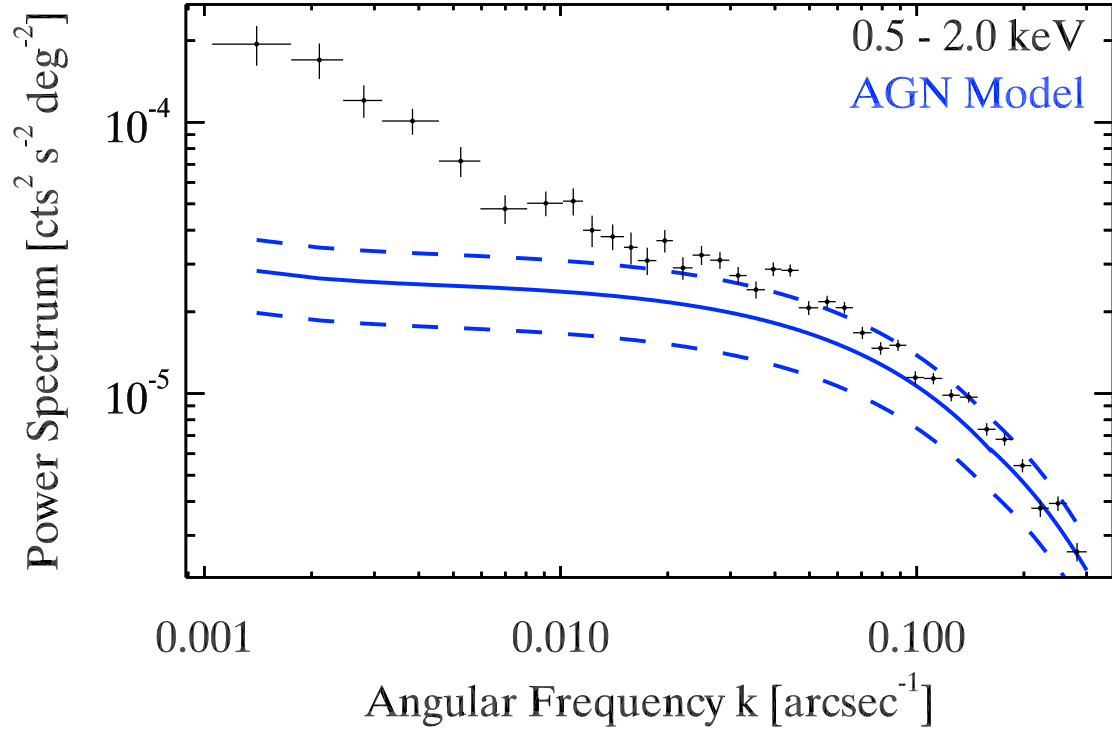


Figure 4.14: The power spectrum of the brightness fluctuations of the unresolved CXB of XBOOTES in the $0.5 - 2.0 \text{ keV}$ band. Blue curve: AGN clustering model (Sect. 4.4.6, including the PSF-smearing model) with a $\pm 30\%$ envelope.

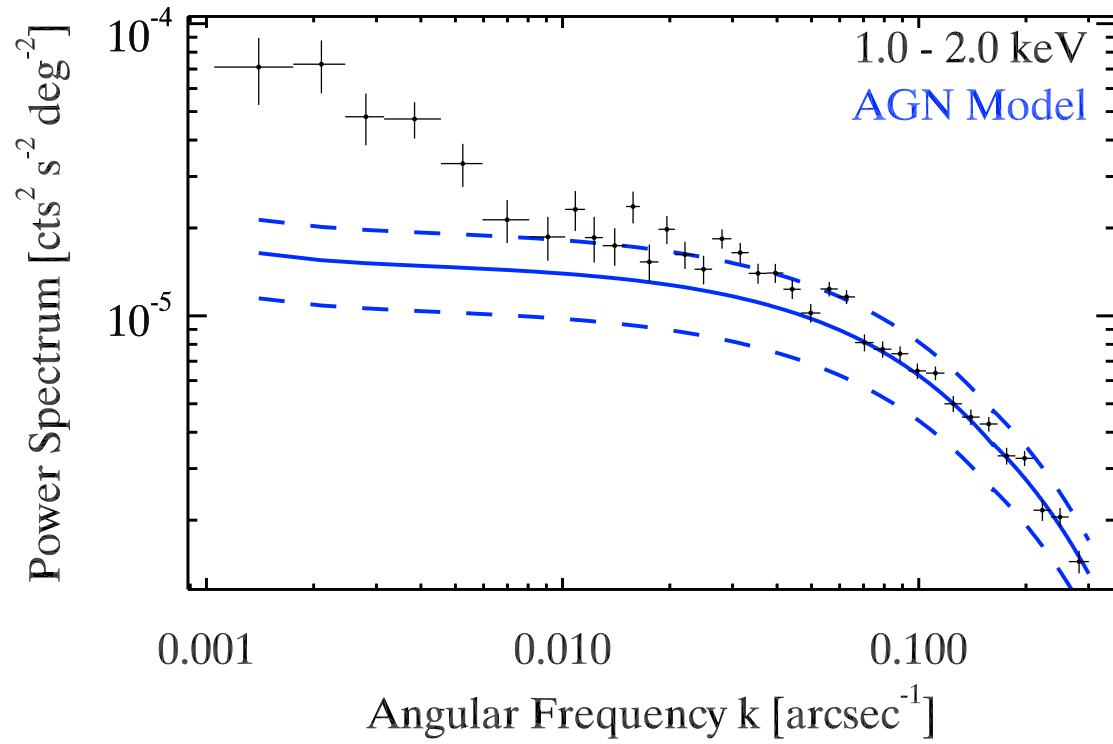


Figure 4.15: The power spectrum of the brightness fluctuations of the unresolved CXB of XBOOTES in the 1.0 – 2.0 keV band. Blue curve: AGN clustering model (Sect. 4.4.6, including the PSF-smearing model) with a $\pm 30\%$ envelope.

that AGN reside at the center of their DMH (e.g. Starikova et al. 2011).

Under this assumption, the AGN one-halo-term becomes equal to the AGN shot-noise term:

$$P_{\text{AGN}}^{(1H)} = \int dS S^2 \frac{dN}{dS} (1 - f(S)) . \quad (4.18)$$

Hereby, dN/dS is the differential $\log N - \log S$ of AGN from L12 and $f(S)$ is the survey selection function of XBOOTES for point sources (Fig. 12 of K05, also see the description for Eq. (4.6)-(4.7) in Sect. 4.3.4.1). We obtain for the one-halo-term a value of $P_{\text{AGN}}^{(1H)} \approx 1.25 \times 10^{-27} (\text{erg cm}^{-2} \text{s}^{-1})^2 \text{deg}^{-2}$. We convert the value into instrumental units with an absorbed powerlaw with a photo-index of $\Gamma = 1.7$ and obtain $P_{\text{AGN}}^{(1H)} \approx 2.49 \times 10^{-5} (\text{cts s}^{-1})^2 \text{deg}^{-2}$. However, using a photo-index between 1.4 and 2.0 results in a variation of $\sim \pm 25\%$ and between 1.6 and 1.8 gives a variation of already $\sim \pm 10\%$. The same holds for the $1.0 - 2.0 \text{ keV}$ band, where we get $P_{\text{AGN}}^{(1H)} \approx 4.67 \times 10^{-28} (\text{erg cm}^{-2} \text{s}^{-1})^2 \text{deg}^{-2} = 1.47 \times 10^{-5} (\text{cts s}^{-1})^2 \text{deg}^{-2}$. Note that this variation comes dominantly from the survey selection function $f(S)$, which is convert from the $0.5 - 7.0 \text{ keV}$ band.

We can write the two-halo-term with Limber's approximation (assuming small angles, $k^{-1} \ll 1 \text{ rad}$) as following:

$$P_{\text{AGN}}^{(2H)}(k) = \int dz \left(\frac{dS}{dz} \right)^2 \left(\frac{d^2 V(z)}{dz d\Omega} \right)^{-1} P_{3D, \text{AGN}} \left(k_{3D} = \frac{\alpha k}{\beta r(z)}, z \right) . \quad (4.19)$$

Hereby, $dS(z)/dz [\text{erg cm}^{-2} \text{s}^{-1} \text{deg}^{-2}]$ is the differential flux distribution as function of redshift, defined in Eq. (4.6) (Sect. 4.3.4.1), $d^2 V(z)/dz/d\Omega [\text{Mpc}^3 \text{h}^{-3} \text{deg}^{-2}]$ is the co-moving volume element, and $r(z) [\text{Mpc h}^{-1}]$ is the co-moving distance to redshift z (e.g. Hogg 1999). The α and β are equal to 2π , if there is a 2π in the exponent of the 2D and 3D Fourier transform, respectively. Otherwise there are equal to one. The AGN 3D power spectrum is computed as following (e.g. Cooray & Sheth 2002):

$$P_{3D, \text{AGN}}(k_{3D}, z) = b(M_{\text{eff}}, z)^2 g(z)^2 P_{\text{lin}}(k_{3D}) . \quad (4.20)$$

Hereby, $P_{\text{lin}}(k_{3D}) [\text{Mpc}^3 \text{h}^{-3}]$ is the 3D linear ΛCDM power spectrum at $z = 0$, computed with the fitting formulae of Eisenstein & Hu (1998), $g(z)$ is the linear growth function (e.g. Dodelson 2003), and $b(M_{\text{eff}}, z)$ is the AGN linear clustering bias factor, computed with the analytical model of Sheth et al. (2001). For the effective mass M_{eff} of the DMH, where the AGN reside, we use $2 \times 10^{13} M_{\odot} \text{h}^{-1}$, which is consistent with recent observations up to $z \sim 3$ (e.g. Allevato et al. 2011; Krumpke et al. 2012; Mountrichas et al. 2013). We should note that in the considered

frequency range $k > 10^{-3} \text{ arcsec}^{-1}$ ($\lesssim 17'$) the one-halo-term dominates over the two-halo-term.

In Fig. 4.15 we show the power spectrum (black crosses) in the $1.0 - 2.0 \text{ keV}$ band together with our AGN clustering model (blue curve), which also takes our PSF-smearing model (Sect. 4.4.2) into account. We can see that our model follows the shape of the power spectrum within an accuracy of $\sim 30\%$ for $\lesssim 2'$ ($\gtrsim 0.008 \text{ arcsec}^{-1}$) but we can see a clear excess of the measured power spectrum in comparison to our AGN clustering model for angular scales of $\gtrsim 2'$. Before we focus on this excess in Sect. 4.4.7, we like to have a look on the angular scales of $\lesssim 2'$, where we can see already some significant deviations. We can see in Fig. 4.15 that in the frequency regime of $[0.8, 7.0] \times 10^{-2} \text{ arcsec}^{-1}$ the power spectrum is by a factor of ~ 1.3 larger than our AGN clustering model. There are several possible explanations for this discrepancy, which are not mutually exclusive.

The most interesting case would be that our definition for the one-halo-term with Eq. (4.18) is too simple and that there is indeed a significant fraction of DMHs, which host more than one AGN. However, it is difficult to have a good quantitative description of the one-halo-term of the unresolved AGN of XBOOTES, with a median redshift of $z \sim 1$ and a median at $L_{0.5-2.0 \text{ keV}} \sim 10^{42.6} \text{ erg s}^{-1}$ (see Sect. 4.3.4.1) with the results of previous AGN clustering studies. These clustering studies use resolved AGN with a luminosity of at least $L_{0.5-2.0 \text{ keV}} \sim 10^{42} \text{ erg s}^{-1}$. Further, the current best measurements of the one-halo-term are performed at low redshift $z < 0.5$, where model parameters from halo occupation distribution (HOD) modeling are still poorly constrained (e.g. Alleinato et al. 2011, 2012, 2014; Krumpe et al. 2010b, 2012; Miyaji et al. 2011; Starikova et al. 2011; Cappelluti et al. 2012a). Those studies nevertheless show for certain that massive DMHs ($\gtrsim 10^{14} h^{-1} M_{\odot}$) can host more than one AGN. Therefore, a simple interpretation of our measured discrepancy could be that up to $\sim 30\%$ the unresolved AGN of XBOOTES reside in a DMH with a second AGN, whereas the rest are single AGN in their DMH. We would like to test this hypothesis in our subsequent study. We can conclude that the measured discrepancy could be explained by our insufficient knowledge of the one-halo-term of faint AGN and that clustering studies with the unresolved CXB have the potential to improve this situation. However, there are still additional factors, which also have to be taken into account.

The higher amplitude of the power spectrum may also indicate that there are more unresolved AGN contributing to the unresolved CXB than known $\log N - \log S$ of AGN suggests. In this respect we compared the $\log N - \log S$ of AGN of XBOOTES (K05) and L12. Since they appear quite consistent in the flux range of the resolved AGN population of XBOOTES, there is no obvious reason why they should

be inconsistent for the unresolved AGN population. Further, given also the good knowledge of the number density of X-ray point sources down to $\sim 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1}$ ($0.5 - 2.0 \text{ keV}$) (e.g. Kim et al. 2007; Georgakakis et al. 2008, L12), it is unlikely that the unresolved AGN population can be significant larger than what the $\log N - \log S$ of L12 predicts. We can nevertheless expect some uncertainty in estimating the one-halo-term due to cosmic variance.

We also should mention that a change of the photo-index of our powerlaw from 1.7 to 1.8 already increases the one-halo-term by $\sim 10\%$. Our spectral fit of the unresolved CXB favors a photo-index of 1.7 but this powerlaw also includes the extragalactic emission from other unresolved sources besides AGN, such as galaxy clusters & groups. A slightly higher photo-index would not be in disagreement with the bulk of point sources in the XBOOTES survey and with AGN sample of much deeper surveys (e.g. Gilli et al. 2007; Yang et al. 2015). However, neither would be a slightly lower one. This adds an additional uncertainty to the estimation of the one-halo-term.

Given these observational uncertainties, one could argue that our AGN clustering model is sufficient to approximately describe the measured power spectrum up to angular scales of $\sim 2'$ ($\gtrsim 0.008 \text{ arcsec}^{-1}$). However, these uncertainties are insufficient to explain the clear excess for angular scales of $\gtrsim 2'$.

4.4.7 An excess at angular scales of $\gtrsim 2'$

We measure a clear excess in our power spectrum in comparison to our AGN clustering model for angular scale of $\gtrsim 2'$. We can exclude that it has instrumental origin, since we show in Fig. 4.12 that the power spectrum of the instrumental background is more than an order of magnitude smaller than our measured power spectrum. We have also tested unsuccessfully if an increase of the effective mass M_{eff} of the DMH, where the AGN reside, up to $10^{14} M_{\odot} \text{ h}^{-1}$ could explain the excess. We can rule out the remaining counts of resolved point sources as an origin, since we show in our test (Fig. 4.2) in Sect. 4.2.2.1 that the fraction of such counts can be neglected.

Since we measure this excess not just in the $0.5 - 2.0 \text{ keV}$ band (Fig. 4.14) but also in the $1.0 - 2.0 \text{ keV}$ band (Fig. 4.15), where the contribution of the unresolved emission of our own Galaxy should be insufficient, it indicates that it is extragalactic in origin. Therefore, we focus first on extragalactic sources.

4.4.7.1 Normal galaxies

In terms of unresolved X-ray point sources, normal galaxies are our best and only candidate. However, given their much smaller number density at the depth of

XBOOTES (e.g. L12, Fig. 5) and a factor of ~ 2 smaller effective clustering bias factor (e.g. Helgason et al. 2014, Fig. 8) in comparison to AGN, we know that their two-halo-term is significantly smaller than for AGN. For the same reason we do not expect that the cross-correlation term between AGN and normal galaxies have a significant amplitude. The same holds for the one-halo- and shot-noise-term but since these are flat components at large scales, they would not be able to explain the measured excess in any case. The shot-noise term of normal galaxies, computed with Eq. 4.18 and the $\log N - \log S$ of L12, would increase our model amplitude by $\sim 5\%$, which is almost insignificant at the given S/N of our measured power spectrum. Hence, we conclude that unresolved normal galaxies are not able to explain the excess.

4.4.7.2 Galaxy clusters & groups and WHIM

In terms of unresolved extended X-ray sources, galaxy clusters & groups are our best and only candidate. The number density and therefore the surface brightness of galaxy clusters & groups below the depth of XBOOTES is not very well known (e.g. Rosati et al. 2002; Finoguenov et al. 2007, 2010; Clerc et al. 2012; Vajgel et al. 2014, see also Table 4.7). When removing the resolved galaxy clusters & groups from the XBOOTES, we are removing the brightest galaxy clusters & groups with fluxes of about $\gtrsim 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$ (K05, also see Sect. 4.2.2). Based on the luminosity-mass scaling relation of Vajgel et al. (2014, Table 6) we can estimate that this includes all galaxy clusters & groups with $\gtrsim 10^{13} M_{\odot}$ and $\gtrsim 3 \times 10^{43} \text{ erg s}^{-1}$ (0.5 – 2.0 keV) for a redshift of $z \lesssim 1$.

Since the clustering signal of galaxy clusters & groups is difficult to describe analytically, clustering studies prefer to use adjusted mock maps of cosmological hydrodynamical simulations in order to compute the theoretical power spectrum of galaxy clusters & groups for a given survey (e.g. Roncarelli et al. 2007; Cappelluti et al. 2012b; Roncarelli et al. 2012; Helgason et al. 2014). In order to test if galaxy clusters & groups are able to explain the excess, we will use the analytical work of Cheng et al. (2004). Their analytical power spectrum (their Fig. 1) is in very good agreement with the measured power spectrum of the simulations of Roncarelli et al. (2007, Fig. 13). Based on the work of Cheng et al. (2004) we know that for angular scales of $\gtrsim 2'$ the galaxy clusters & groups power spectrum is dominated by low-redshift galaxy clusters & groups of $z < 0.1$ (their Fig.1b, $\ell \lesssim 5 \times 10^3$) and its amplitude is reduced by more than two orders of magnitude after removing all galaxy clusters & groups with masses of $> 10^{13} M_{\odot}$ (their Fig.1c), which should have been all detected in XBOOTES as resolved sources.

In the following, we will use the theoretical power spectrum of Cheng et al. (2004) for the mass group $< 10^{13} M_{\odot}$ (their Fig.1c, dashed curve) as our galaxy clusters & groups clustering model, which presents an upper limit for the XBOOTES survey. Since the amplitude of this power spectrum is directly correlated with the surface brightness of the galaxy clusters & groups, we can use our estimation for XBOOTES (Table 4.7) as constraints for the amplitude. To have an upper limit we use the surface brightness estimate of $5 \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ deg}^{-2}$ (0.5 – 2.0 keV) based on the $\log N - \log S$ of Rosati et al. (2002), which is also consistent with more recent measured $\log N - \log S$ of Finoguenov et al. (2007, 2010). In order to convert the galaxy clusters & groups power spectrum into instrumental units and from the 0.5 – 2.0 keV to the 1.0 – 2.0 keV band, we assume for the average spectral model of the unresolved galaxy clusters & groups an one-temperature model with an absorbed APEC of $T = 1.0$ keV. This is a typical temperature for galaxy clusters & groups of $\sim 10^{13} M_{\odot}$, using the mass-temperature scaling relation of Sun et al. (2009, Table 6). By adding this clustering model to our AGN clustering model, we can not increase our model amplitude significantly. We would need to increase the amplitude of the galaxy clusters & groups power spectrum by about two orders of magnitude or the surface brightness by a factor of ~ 10 in order to match the excess.

Some circumstances in principle are able to reduce this discrepancy. For instance, using a higher temperature in our spectral model, we can increase the amplitude up to a factor of ~ 2.5 in the 1.0 – 2.0 keV band. However, since the unresolved galaxy clusters & groups will be dominated by $< 10^{13} M_{\odot}$ objects a lower temperature than $T = 1.0$ keV appears more reasonable, under the assumption the unresolved galaxy clusters & groups follow the mass-temperature scaling relations of Sun et al. (2009, Table 6). Further, there is also a model uncertainty of the power spectrum of Cheng et al. (2004, Fig. 3) due different implementations of non-gravitational effects, which change the properties and distribution of the ICM in the galaxy clusters & groups. However, this uncertainties can not explain the large discrepancy between the power spectrum of galaxy clusters & groups and the measured excess. Therefore, we also conclude that unresolved galaxy clusters & groups are not able to explain the excess.

The recent cosmological hydrodynamical simulations of Roncarelli et al. (2012, Fig. 6) and Ursino et al. (2014, Fig. 3) predict that the clustering signal of WHIM is between one and two orders of magnitude smaller than that of galaxy clusters & groups. When removing the resolved galaxy clusters & groups in our survey, we also inevitably remove a significant fraction of the corresponding WHIM in our observations. This will further decrease the clustering signal. If we trust the definition of WHIM, as it was used in these simulations, then we expect that they neither are able explain the excess.

Since the unresolved galaxy clusters & groups and WHIM are expected to be rather close-by whereas the unresolved AGN are at high redshift and also given the smaller number density, we do not expect that their cross-correlation term with AGN have a significant amplitude.

4.4.7.3 Energy spectrum of the excess

Based on the current knowledge of the clustering properties of known source populations in X-ray surveys, we are not able to explain the excess. In order to gain some more constraints on the origin of the excess, we analyze the energy spectrum of the emission, which causes the excess.

For constructing the energy spectrum of the excess, we compute the following quantity for each energy bin:

$$\langle P_{\text{Excess}} \rangle = \frac{\sum_{k_1}^{k_2} P(k)}{\sum_{k_1}^{k_2}} - \frac{\sum_{k_3}^{k_4} P(k)}{\sum_{k_3}^{k_4} P_{\text{PSF}}(k)} \frac{\sum_{k_1}^{k_2} P_{\text{PSF}}(k)}{\sum_{k_1}^{k_2}}, \quad (4.21)$$

which subtracts the shot-noise term of known source populations, mainly AGN, from the power spectrum. Hereby, $[k_1, k_2]$ stands for the frequency interval $[0.2, 0.6[\times 10^{-2} \text{ arcsec}^{-1}$, where the excess dominates the power spectrum, and $[k_3, k_4]$ for the frequency interval $[4, 10[\times 10^{-2} \text{ arcsec}^{-1}$, where the shot-noise term of known source populations dominate the power spectrum, and $P_{\text{PSF}}(k)$ is our PSF-smearing model (Sect. 4.4.2).

The result is shown in Fig. 4.16 (black crosses). We detected a clear signal in all energy bins. In panel (a) we also show the energy dependence of the power spectrum (gray crosses) before subtracting the extrapolated power spectrum amplitude of the frequency interval $[k_3, k_4]$. We can see that the spectral shape does not change very significantly, which makes the subtraction of the average shot-noise term of the known source populations in Eq. (4.21) not very important. Further, it shows that the shot-noise term of known source populations has an almost insignificant contribution to the power spectrum for angular scales of $\gtrsim 3'$ ($\lesssim 0.006 \text{ arcsec}^{-1}$).

To look for dependences due to our choice of the frequency interval for averaging the power spectrum, we also tested smaller and larger frequency intervals (up to 0.02 arcsec^{-1}). This results mostly in an overall decrease of the S/N but the spectral shape remains very similar. The same is also true for larger or smaller exclusion areas for the resolved point sources (Sect. 4.2.2).

We should be aware that the first energy bin ($0.5 - 0.7 \text{ keV}$) of the spectrum has the highest potential to be contaminated by brightness fluctuations caused by Galactic emission (also see Sect. 4.4.5), based on our spectral fit of the unresolved

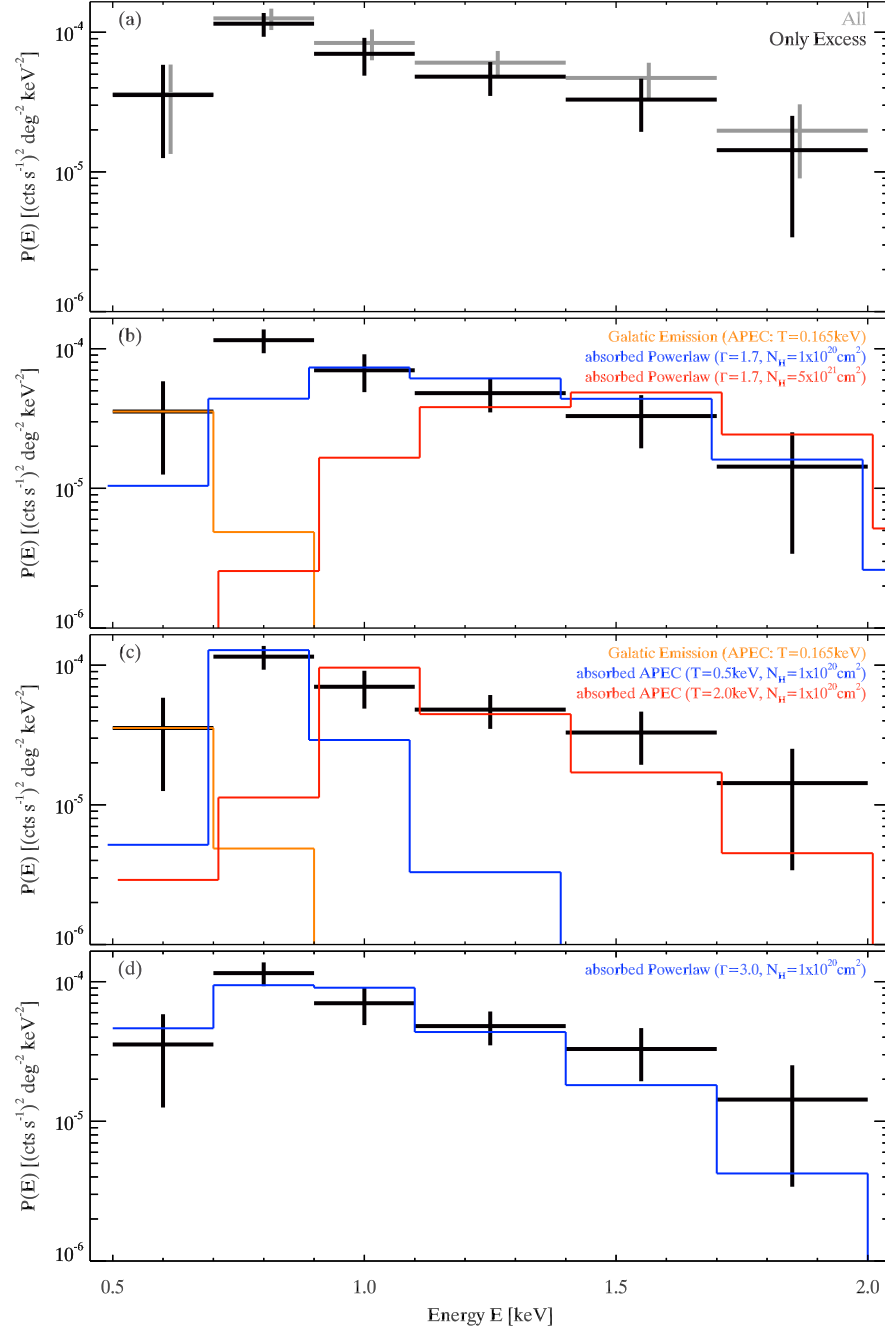


Figure 4.16: Average power spectrum of the excess (black crosses) as a function of energy computed with Eq. (4.21) together with several normalized spectral models. The gray crosses in the upper panel show the average power spectrum for the frequency interval $[0.2, 0.6] \times 10^{-2} \text{ arcsec}^{-1}$ (first term in Eq. 4.21).

CXB (Sect. 4.3.2). Therefore, we include our spectral model of the Galactic emission (APEC model of $T = 0.164$ keV) of this spectral fit also in Fig. 4.16 (orange lines in panel (b)-(c)). From this, we can see that for the adjacent energy bins (> 0.7 keV) we can neglect the contribution of the Galactic emission.

In principle, Galactic emission, which may cause measurable brightness fluctuations, does not need to have the same spectral shape as we measure in our spectral fit of the unresolved CXB of XBOOTES. However, from the work of Henley & Shelton (2013) we know that the Galactic halo emission, the hottest component of the Galactic emission, can not be much hotter than ~ 0.250 keV and that it has an average value of ~ 0.204 keV for the upper quartile of the norther Galactic hemisphere (their Table 2), where XBOOTES is located. If we would use a value of $T = 0.204$ keV for the APEC model instead, the spectrum of the Galactic emission would not change significantly at the given energy resolution. Hence, we conclude that we can neglect any contamination of brightness fluctuations caused by Galactic emission for > 0.7 keV but probably not for the energy bin of $0.5 - 0.7$ keV. Since we will see in the following (panel (b)-(c)) that none of the spectral models of known extragalactic source populations is able to explain the clustering signal in the energy bin of $0.5 - 0.7$ keV, this may suggests that this energy bin is indeed contaminated by a clustering signal of Galactic emission.

Due to the complicated procedure involved in producing the energy spectrum in Fig. 4.16, we can not conduct a rigorous spectral fitting. For this reason in the current work, we will compare the measured spectrum with typical one-component spectra of known source populations in panel (b) and (c).

In panel (b) we show two absorbed powerlaws with a photo-index of $\Gamma = 1.7$ but different absorption. These are typical spectra of X-ray point sources, such as AGN and normal galaxies, for XBOOTES (K05) and most other X-ray surveys (e.g. Reeves & Turner 2000; Gilli et al. 2007; Georgakakis et al. 2008; Corral et al. 2011; Reynolds et al. 2014; Ueda et al. 2014; Yang et al. 2015). We can see that the powerlaw with the smallest absorption ($N_{\text{H}} = 10^{20} \text{ cm}^{-2}$, blues lines) is only able to explain the spectrum for $\gtrsim 0.9$ keV. The situation is even worse for larger absorption (e.g. $N_{\text{H}} = 5 \times 10^{21} \text{ cm}^{-2}$, red lines). We can also see that a linear combination of both models would not improve the situation. Based on this we can conclude that typical X-ray point sources are not able to explain the entire spectrum of the excess. We note that even for known extremes cases, where AGN have a steep powerlaw with a photo-index of up to $\Gamma = 2.5$, we come to the same conclusion (also see description for panel (d)).

In panel (c) we show two absorbed APEC models with the temperatures of $T = 0.5$ keV and $T = 2.0$ keV. They represent typical spectra for galaxy groups and

galaxy clusters (e.g. Sun et al. 2009; Sun 2012; Vajgel et al. 2014; Böhringer et al. 2014). We can see that both spectra are not able to explain entire spectrum of the excess on its own. Here we should also take into account that we expect that the unresolved galaxy clusters & groups are not much hotter than ~ 1.0 keV, under the assumption that they follow the mass-temperature scaling relations of Sun et al. (2009, Table 6) (also see Sect. 4.4.7.2). With the current definition of WHIM, they are also expected to not be hotter than ~ 1.0 keV (e.g. Roncarelli et al. 2012; Ursino et al. 2014). Hence, we can conclude that neither galaxy clusters & groups nor WHIM are able to explain the entire spectrum of the excess.

The complexity of the spectrum may indicate that the excess is caused by more than one type of source. For example, one could think of a linear combination of different models (e.g. APEC of $T = 0.5$ keV plus powerlaw of $\Gamma = 1.7$) to explain the entire spectrum of the excess. We plan to perform a more detailed analysis of this type in the future.

In panel (d) we also show that in principle an absorbed powerlaw with a photo-index of $\Gamma = 3.0$ and $N_{\text{H}} = 10^{20} \text{ cm}^{-2}$ is able to explain the entire spectrum. However, it is difficult to associated any known extragalactic source population with such a spectral shape.

At this point, we can not draw any strong conclusions from the spectrum of the excess. In summary, it is not possible to describe the entire spectrum with a typical spectrum of any known extragalactic source population and the complexity of the spectrum may indicate that the excess is produces by more than one type of source. In our future study we will investigate, if the excess also dominants the power spectrum at larger angular scales ($> 17'$).

4.5 Discussion and summary

Studies of brightness fluctuations of the unresolved CXB are a great opportunity to study clustering properties of faint source populations, which are otherwise difficult to obtain with clustering studies of the resolved CXB due to lack of deep and wide surveys and selection effects, such as S/N-cut for spectroscopic redshift or luminosity-cut for AGN identification. By using the currently largest continuous *Chandra* survey, XBOOTES, we perform the currently most accurate measurement of the brightness fluctuations of the CXB for angular scales of $\lesssim 17'$. We find that known source populations are only able to describe these fluctuations for angular scales smaller than $\sim 2'$.

As a preparation for measuring the power spectrum of the brightness fluctuations we performed following tasks to ensure a clean measurement:

- We made a spectral fit of the unresolved CXB emission of XBOOTES in order to disentangle its components and show that the surface brightness of the Galactic and extragalactic component are consistent with previous observations (Sect. 4.3).
- We have determined for the first time the average power spectrum model of the PSF-smearing for the entire FOV of *Chandra*'s ACIS-I with $\sim 5\%$ accuracy. We find that with its help we are able to describe well the measured power spectrum down to very small angular scales (Sect. 4.4.2).
- We have evaluated different methods to estimate the photon shot-noise component of the power spectrum and show that our default estimator works well (Sect. 4.4.3).
- We have measured for the first time the brightness fluctuations of the instrumental background of *Chandra*'s ACIS-I and show that their contribution is neglectable for the measured power spectrum in the energy range $0.5 - 2.0$ keV (Sect. 4.4.4.3).
- We show that the mask effect does not change the measured power spectrum very significantly (Sect. 4.4.4.1).
- We show that for our fluctuation analysis we do not have to subtract the instrumental background from the observation (Sect. 4.4.4.2) and that we can neglect the effect of vignetting at the given S/N level (Sect. 4.4.4.4).

4.5.1 Unresolved AGN

We find that at the depth of XBOOTES (average point-source sensitivity: $S_{0.5-2.0\text{ keV}} \sim 2.3 \times 10^{-15} \text{ erg cm}^{-2} \text{ s}^{-1}$) AGN are the major extragalactic component of the unresolved CXB, accounting for $\sim 30\%$ of the extragalactic emission (Table 4.7). They represent an AGN population at high redshift (median $z \sim 1.0$) and low luminosity (median $L_{0.5-2.0\text{ keV}} \sim 10^{42.6} \text{ erg s}^{-1}$, Sect. 4.3.4.1). The clustering properties of these AGN are currently extrapolated from clustering studies of more luminous and more nearby resolved AGN populations.

We also find that this AGN population dominates the brightness fluctuations of the unresolved CXB of XBOOTES up to angular scales of $2'$ and that a simple AGN clustering model can describe the power spectrum of these fluctuations reasonably well, although not perfectly (Sect. 4.4.6). We see a discrepancy of $\sim 30\%$ for angular scales of $\sim 0.2' - \sim 2'$. In this regime, our AGN clustering model is defined by the

clustering properties of AGN within their DMH (one-halo-term), which is in general not very well constrained even for the resolved AGN population. The detected discrepancy indicates that not all AGN reside alone in one DMH. In our subsequent study, we plan to investigate this with a more sophisticated model. It represents a great opportunity to improve our knowledge of the clustering properties of AGN on small spatial scales ($\lesssim 2 h^{-1}$ Mpc) at high redshift ($z \sim 1$).

In the same subsequent study, we plan to expand our measurement to larger scales ($> 17'$) up to 3° , which is the maximum possible angular scale of XBOOTES. With this range of angular scales, we will also be able to test our understanding of the clustering properties of AGN between different DMHs (two-halo-term) and to measure the average DMH mass of unresolved AGN. However, in order to expand to these scales we have to create a mosaic of all XBOOTES observation, which introduces further complexities to our study.

4.5.2 The excess in the clustering signal

We are able to describe the brightness fluctuations with our AGN clustering model up to angular scales of $2'$ with $\sim 30\%$ accuracy. For larger scales we measure a clear excess in comparison to this model, which we can not explain with any known source population (Sect. 4.4.7). We have tested several possibilities with negative results:

- Instrumental background (Sect. 4.4.4.3).
- Remaining counts of resolved point sources (Sect. 4.2.2.1).
- Increasing the DMH mass in our AGN clustering model (Sect. 4.4.6) up to $10^{14} M_\odot h^{-1}$.
- Normal Galaxies (Sect. 4.4.7.1).
- Galaxy clusters & groups and WHIM (Sect. 4.4.7.2).

Since we measure the excess not only in the $0.5 - 2.0$ keV but also in the $1.0 - 2.0$ keV band, an extragalactic origin appears to be more likely or at least to be the dominant source.

We have measured the spectral shape of the excess (Sect. 4.4.7.3), which unfortunately did not lead to any definitive conclusions either. In accord with the above conclusions the energy spectrum can not be described by a typical spectrum of known X-ray source populations (AGN, normal galaxies, galaxy clusters & groups or WHIM). The complexity of the spectrum may be an indication that more than

one type of source causes the excess. However, we also find that the energy spectrum can be described with a very steep powerlaw, which can not be associated with any known source population. For the $0.5 - 0.7$ keV band, we can not exclude the possibility that Galactic emission makes a significant contribution to the observed power spectrum.

4.5.3 eROSITA forecast

The all-sky survey of eROSITA (eRASS) has great potential to perform LSS studies with the resolved part of the CXB (see Chapter 3 for details) but also with the unresolved part, as we show here. The survey covers an extragalactic sky of $\sim 34\,000$ deg² (excluding the Galactic plane, $|b| > 10^\circ$), has an average point-source sensitivity of $S_{0.5-2.0\text{ keV}} \approx 1.1 \times 10^{-14}$ erg cm⁻² s⁻¹, and will detect about 3 million AGN (number density ~ 84 deg⁻²) (see Table 2.2 in Chapter 2 for details). About $21\,000$ deg² ($\sim 60\%$) of the survey will consist of unresolved emission, assuming for resolved point sources an average circular exclusion area with a radius of $140''$, which is more than four times the size of the survey-averaged PSF ($\sim 30''$) and about two times the size of the PSF at the largest offset angle ($\sim 70''$) (Merloni et al. 2012, Fig. 2.1.3). With such a large sky coverage we will be able to dramatically improve our measurement of the brightness fluctuations of the unresolved CXB in terms of S/N by up to a factor of ~ 50 as compared to XBOOTES. Further, it will cover all required angular scales, which we need to study the clustering properties of unresolved source populations.

The average point-sources sensitivity of eRASS is about five times shallower than XBOOTES. However, the unresolved AGN will still represent an interesting AGN population. We show their predicted differential flux distribution in Fig. 4.17, using Eq. (4.6)-(4.7) from Sect. 4.3.4.1. It shows that this AGN population will be at high redshift (median $z \sim 1.2$, peak $z \sim 1.1$) and will have a moderate luminosity (median $L_{0.5-2.0\text{ keV}} \sim 10^{43.4}$ erg s⁻¹, peak $\sim 10^{43.8}$ erg s⁻¹). With eRASS we will be able to measure their clustering properties to high accuracy. It enables us to tell to how many AGN reside in average in a DMH and what is the fraction of AGN as satellite source within one DMH, which are important questions for AGN evolution and its implications. These properties can be derived from the power spectrum of the brightness fluctuations for angular scales of $\lesssim 2'$. For larger angular scales, we will be able to further investigate the origin of the excess found in this work, and to study the two-halo-term of AGN in more detail.

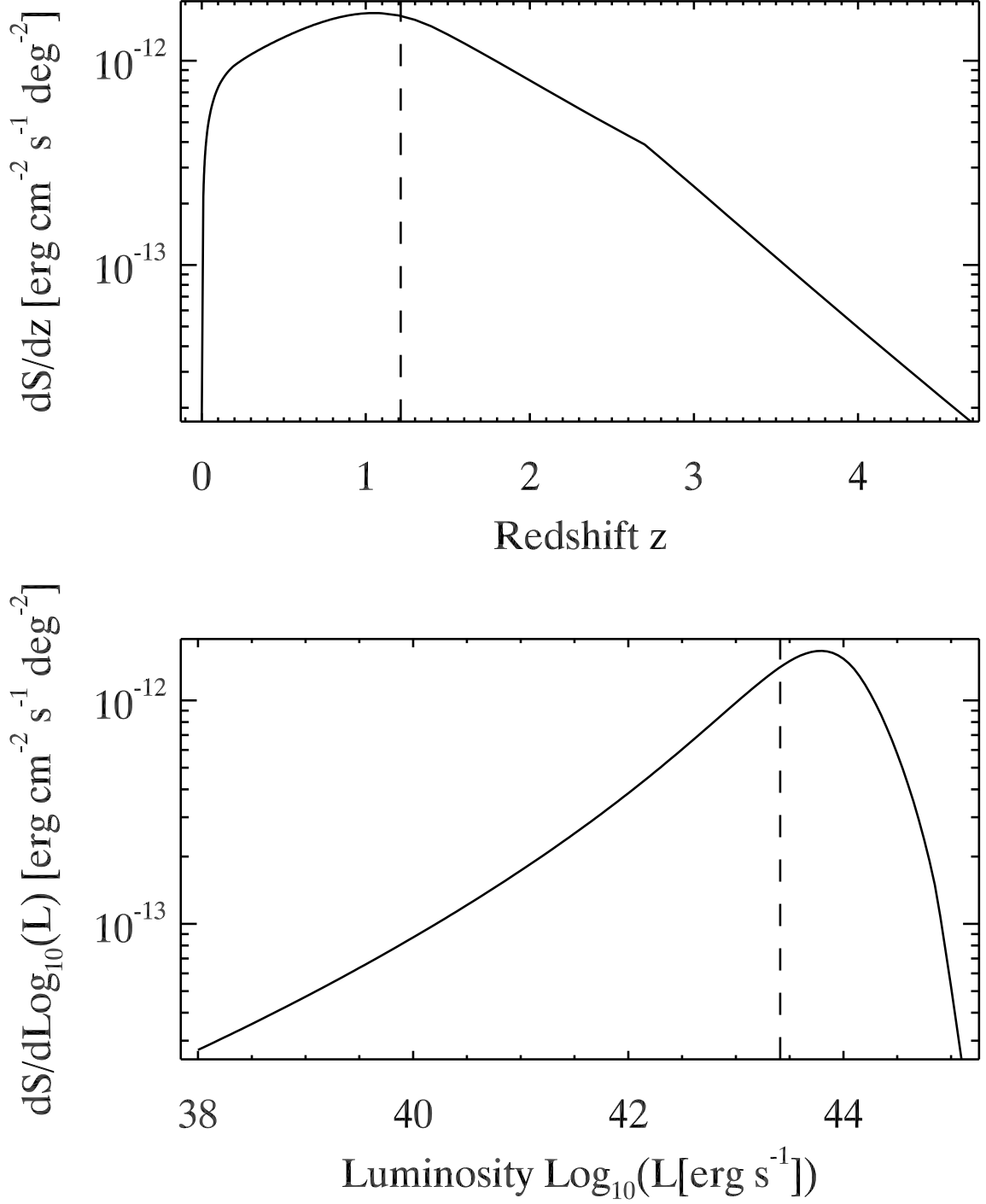


Figure 4.17: Predicted differential flux distribution (alias the flux production rate per solid angle, Eq. 4.6-4.7) for the unresolved AGN in the eROSITA all-sky survey in the $0.5 - 2.0 \text{ keV}$ band as a function of redshift (*top*) and luminosity (*bottom*), assuming an average point-source sensitivity of $S_{0.5-2.0 \text{ keV}} = 1.1 \times 10^{-14} \text{ erg cm}^{-2} \text{ s}^{-1}$. The dashed lines show the median.

Conclusions

Large-scale structure (LSS) studies with X-ray surveys are a powerful tool to improve our understanding of the growth of supermassive black holes (SMBHs) over cosmic time and its implications for galaxy evolution, and to improve the constraints of the standard model of cosmology. In this work, we demonstrate that the scientific significance of LSS studies using active galactic nuclei (AGN) in X-ray surveys can be greatly enhanced by performing wider and deeper X-ray surveys, and by using the unresolved source populations of X-ray surveys. In respect to the former, we have studied the prospects of using the AGN sample to be detected by the upcoming eROSITA all-sky survey (eRASS) for LSS studies. In respect to the latter, we have conducted a LSS study with the unresolved cosmic X-ray background (CXB) by analyzing its surface brightness fluctuations with the XBOOTES survey conducted by the *Chandra* telescope.

Large-scale structure studies with AGN - prospects for eRosita

We have studied the statistical properties of AGN to be detected in eRASS. We show that eROSITA will detect about 3 million AGN in its four-year all-sky survey in the $0.5 - 2.0$ keV band, with a point-source sensitivity of $\sim 1.1 \times 10^{-14}$ erg s $^{-1}$ cm $^{-2}$. This will be by far the largest sample of X-ray selected AGN. We show that this survey will detect about 30 times more AGN and will be about 30 times better in sensitivity than its ~ 25 year old predecessor, the ROSAT all-sky survey (RASS). The AGN sample of eRASS will have a median redshift of $z \approx 1$ and approximately 40% of the AGN will be located in the redshift range of $z = 1 - 2$, where the bulk of the X-ray emission of unobscured AGN is produced. eROSITA will also detect for the first time a statistically significant number of AGN at very high redshift, where we still have no good understanding of their properties. Based on the extrapolation of current X-ray luminosity functions (XLFs) of AGN, we predict that about $10^4 - 10^5$ AGN will be detected beyond redshift $z = 3$ and about $2000 - 30\,000$ beyond $z = 4$, which will potentially include some of the earliest AGN in the Universe. The AGN

sample of eRASS will have a median luminosity of $\sim 10^{44} \text{ erg s}^{-1}$, which represents the typical luminosity of the entire population of unobscured AGN. Approximately 10 % of the AGN will have luminosities higher than $\sim 10^{45} \text{ erg s}^{-1}$, improving the current statistics in this luminosity-regime by more than an order of magnitude.

Besides the 0.5 – 2.0 keV band, eROSITA will detect in the 2.0 – 10.0 keV band about 10^5 AGN with a minimum flux of $\sim 1.8 \times 10^{-13} \text{ erg s}^{-1} \text{ cm}^{-2}$. These AGN will have a median redshift of $z \approx 0.4$ and a typical luminosity of $\sim 10^{44.4} \text{ erg s}^{-1}$. This will be the first all-sky survey ever in this energy band, improving our picture of strongly obscured AGN in the 0.5 – 2.0 keV band dramatically.

We discussed how the unique statistical properties of both AGN samples will revolutionize our understanding of the entire AGN population and its implications for galaxy evolution. We demonstrate that the eRASS-AGN sample will be able to significantly improve our current knowledge of the AGN spatial density as a function of redshift and luminosity over a wide range of cosmic time. This will allow us to study the growth and evolution of SMBHs to an unprecedented accuracy.

We show that the detected AGN in eRASS will also supply us with outstanding opportunities for detailed LSS research. It will be possible, for the first time, to perform detailed redshift- and luminosity-resolved studies of the clustering strength of X-ray selected AGN, using data from a sky area of $\gtrsim 10^4 \text{ deg}^2$. Even with data from sky patches of only $\sim 2500 \text{ deg}^2$, we will be able to measure the redshift evolution of the clustering strength with an accuracy better than $\sim 10\%$. The eRASS-AGN sample will also enable us to expand the luminosity range of AGN clustering studies beyond $L_{0.5-2.0 \text{ keV}} \sim 10^{44} \text{ erg s}^{-1}$. This will make statistically meaningful comparison of the clustering properties of luminous X-ray selected AGN and optical selected AGN (quasars) possible for the first time. Note that AGN clustering studies with eRASS will yield meaningful results long before the full four-year survey will be completed. All these measurements will dramatically improve our understanding of AGN triggering mechanisms and its dependence on the AGN environment, such as the host galaxy and dark matter halo (DMH) properties. This brings us a huge step closer to answering the questions of when and how SMBHs grow and how they co-evolve with their DMH over cosmic time. These are essential questions for understanding galaxy evolution.

We demonstrate for the first time that, given the breadth and depth of eRASS, we will be able to use its AGN sample as a cosmological probe. It will be possible for the first time to convincingly detect baryon acoustic oscillations (BAOs) at high statistical significance ($\sim 10\sigma$) with AGN clustering measurements. More importantly, it will enable us to make BAO measurements in the currently uncharted redshift regime of $z \sim 1 - 2$, improving significantly our constraints on dark energy.

Finally, we investigated broad requirements for follow-up missions and show that in order to fully exploit the potential of the eRASS-AGN sample, we need redshift information for a large sky fraction. For studies of the spatial density and the clustering strength of AGN, a redshift survey with an accuracy of the order of $\delta z \sim 0.1$, a sky coverage of $\sim 10^4 \text{ deg}^2$, and an optical depth of $I \gtrsim 22.5 \text{ mag}$ will be sufficient to achieve the major goals. The BAO measurement is more demanding. A detection at a 4σ confidence level in the redshift bin of $z = 0.8 - 1.2$ requires a spectroscopic redshift survey with an accuracy of $\delta z \gtrsim 0.01$, a sky coverage of at least $\sim 2 \times 10^4 \text{ deg}^2$, and an optical depth of $I > 22.5 \text{ mag}$.

Large-scale structure studies with the unresolved cosmic X-ray background of XBOOTES

Analysis of the surface brightness fluctuations of the unresolved CXB have a great potential to study the clustering properties of source populations, which are otherwise inaccessible with clustering studies of resolved sources of current X-ray surveys. We have conducted the most accurate measurement to date of the brightness fluctuations of the unresolved CXB for angular scales of $\lesssim 17'$. This was accomplished by using the XBOOTES survey, the currently largest continuous survey of the *Chandra* telescope. We find that the power spectrum of the brightness fluctuations can be described with a conventional AGN clustering model to an accuracy of 30 % for angular scales below $\sim 2'$. Above $\sim 2'$ we measure a significant excess of unknown origin with up to an order of magnitude difference in comparison to the AGN clustering model.

The unresolved AGN population in XBOOTES accounts for 30 % of the unresolved extragalactic emission, is located at high redshift (median $z \sim 1.0$), and has a low luminosity (median $L_{0.5-2.0 \text{ keV}} \sim 10^{42.6} \text{ erg s}^{-1}$). The detected 30 % discrepancy between our measurement and a conventional AGN clustering model for angular scales of $\sim 0.2' - \sim 2'$ suggests that not all AGN reside alone in one DMH as extrapolated from clustering studies of more luminous and more nearby resolved AGN populations. The discrepancy also presents a great opportunity to improve our knowledge of the clustering properties of this unresolved AGN population. We will investigate this in a future study with more sophisticated AGN clustering models.

The measured excess in the power spectrum for angular scales above $\sim 2'$ can not be explained with any single known X-ray source population. An extragalactic origin appears to be most likely or at least to be the dominant source. The complex shape of the energy spectrum of the excess can not be described with a typical spectrum of known X-ray source populations, which may indicate that the excess is caused by

more than one type of source. Finally, we make predictions on how eRASS will be able to advance our understanding of various components of the yet unresolved part of the CXB.

References

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *ApJS*, 182, 543
- Aird, J., Nandra, K., Laird, E. S., et al. 2010, *MNRAS*, 401, 2531
- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ArXiv e-prints*, 1501.00963
- Alexander, D. M. & Hickox, R. C. 2012, *New A Rev.*, 56, 93
- Alexander, D. M., Stern, D., Del Moro, A., et al. 2013, *ApJ*, 773, 125
- Allen, S. W., Evrard, A. E., & Mantz, A. B. 2011, *ARA&A*, 49, 409
- Allevato, V., Finoguenov, A., Cappelluti, N., et al. 2011, *ApJ*, 736, 99
- Allevato, V., Finoguenov, A., Civano, F., et al. 2014, *ApJ*, 796, 4
- Allevato, V., Finoguenov, A., Hasinger, G., et al. 2012, *ApJ*, 758, 47
- Anders, E. & Grevesse, N. 1989, *Geochim. Cosmochim. Acta*, 53, 197
- Anderson, L., Aubourg, E., Bailey, S., et al. 2012, *MNRAS*, 427, 3435
- Anderson, M. E., Churazov, E., & Bregman, J. N. 2014, *ArXiv e-prints*, 1408.4115
- Antonucci, R. 1993, *ARA&A*, 31, 473
- Arnaud, K. A. 1996, in *Astronomical Society of the Pacific Conference Series*, Vol. 101, *Astronomical Data Analysis Software and Systems V*, ed. G. H. Jacoby & J. Barnes, 17
- Baganoff, F. 1999, *ACIS On-orbit Background Rates and Spectra from Chandra OAC Phase 1*, ACIS Memo 162, Massachusetts Institute of Technology, Center for Space Research

- Barcons, X. & Fabian, A. C. 1988, MNRAS, 230, 189
- Bassett, B. & Hlozek, R. 2010, Baryon acoustic oscillations, ed. P. Ruiz-Lapuente, 246
- Blanton, M. R. & Roweis, S. 2007, AJ, 133, 734
- Böhringer, H., Chon, G., & Collins, C. A. 2014, A&A, 570, A31
- Bradt, H. V. D., Ohashi, T., & Pounds, K. A. 1992, ARA&A, 30, 391
- Brand, K., Brown, M. J. I., Dey, A., et al. 2006, ApJ, 641, 140
- Brandt, W. N. & Hasinger, G. 2005, ARA&A, 43, 827
- Brusa, M., Civano, F., Comastri, A., et al. 2010, ApJ, 716, 348
- Brusa, M., Comastri, A., Gilli, R., et al. 2009, ApJ, 693, 8
- Brusa, M., Gilli, R., Civano, F., et al. 2011, Memorie della Societa Astronomica Italiana Supplementi, 17, 106
- Bulbul, E., Markevitch, M., Foster, A., et al. 2014, ApJ, 789, 13
- Burlon, D., Ajello, M., Greiner, J., et al. 2011, ApJ, 728, 58
- Busca, N. G., Delubac, T., Rich, J., et al. 2013, A&A, 552, A96
- Capak, P., Aussel, H., Ajiki, M., et al. 2007, ApJS, 172, 99
- Cappelluti, N., Allevato, V., & Finoguenov, A. 2012a, Advances in Astronomy, 2012, 1
- Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, A&A, 497, 635
- Cappelluti, N., Kashlinsky, A., Arendt, R. G., et al. 2013, ApJ, 769, 68
- Cappelluti, N., Ranalli, P., Roncarelli, M., et al. 2012b, MNRAS, 427, 651
- Carter, C., Karovska, M., Jerius, D., Glotfelty, K., & Beikman, S. 2003, in Astronomical Society of the Pacific Conference Series, Vol. 295, Astronomical Data Analysis Software and Systems XII, ed. H. E. Payne, R. I. Jedrzejewski, & R. N. Hook, 477
- Chambers, K. & the Pan-STARRS team. 2006, in The Advanced Maui Optical and Space Surveillance Technologies Conference

- Chaudhary, P., Brusa, M., Hasinger, G., Merloni, A., & Comastri, A. 2010, *A&A*, 518, A58
- Chaudhary, P., Brusa, M., Hasinger, G., et al. 2012, *A&A*, 537, A6
- Cheng, L.-M., Wu, X.-P., & Cooray, A. 2004, *A&A*, 413, 65
- Chiappetti, L., Clerc, N., Pacaud, F., et al. 2012, *ArXiv e-prints*, 1211.4492
- Civano, F., Brusa, M., Comastri, A., et al. 2011, *ApJ*, 741, 91
- Clerc, N., Sadibekova, T., Pierre, M., et al. 2012, *MNRAS*, 423, 3561
- Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, *MNRAS*, 362, 505
- Colless, M., Dalton, G., Maddox, S., et al. 2001, *MNRAS*, 328, 1039
- Comastri, A., Setti, G., Zamorani, G., & Hasinger, G. 1995, *A&A*, 296, 1
- Comparat, J., Kneib, J.-P., Escoffier, S., et al. 2013, *MNRAS*, 428, 1498
- Cooray, A. & Sheth, R. 2002, *Phys. Rep.*, 372, 1
- Corral, A., Della Ceca, R., Caccianiga, A., et al. 2011, *A&A*, 530, A42
- Corral, A., Page, M. J., Carrera, F. J., et al. 2008, *A&A*, 492, 71
- Dalton, G., Trager, S. C., Abrams, D. C., et al. 2012, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 8446, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, 145, 10
- de Jong, R. S., Bellido-Tirado, O., Chiappini, C., et al. 2012, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 8446, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*
- De Luca, A. & Molendi, S. 2004, *A&A*, 419, 837
- Diego, J. M., Sliwa, W., Silk, J., & Barcons, X. 2003, *MNRAS*, 344, 951
- Dodelson, S. 2003, *Modern cosmology*, 1st edn. (Academic Press)
- Draper, A. R. & Ballantyne, D. R. 2012, *ApJ*, 751, 72

- Ebrero, J., Carrera, F. J., Page, M. J., et al. 2009, *A&A*, 493, 55
- Eisenstein, D. J. & Hu, W. 1998, *ApJ*, 496, 605
- Eisenstein, D. J., Seo, H.-J., Sirko, E., & Spergel, D. N. 2007, *ApJ*, 664, 675
- Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, *ApJ*, 633, 560
- Elvis, M., Civano, F., Vignali, C., et al. 2009, *ApJS*, 184, 158
- Esquej, P., Page, M., Carrera, F. J., et al. 2013, *A&A*, 557, A123
- Fabian, A. C. & Barcons, X. 1992, *ARA&A*, 30, 429
- Fanidakis, N., Baugh, C. M., Benson, A. J., et al. 2012, *MNRAS*, 419, 2797
- Fanidakis, N., Georgakakis, A., Mountrichas, G., et al. 2013, *MNRAS*, 435, 679
- Ferrarese, L. & Ford, H. 2005, *Space Sci. Rev.*, 116, 523
- Ferrarese, L. & Merritt, D. 2000, *ApJ*, 539, L9
- Finoguenov, A., Guzzo, L., Hasinger, G., et al. 2007, *ApJS*, 172, 182
- Finoguenov, A., Watson, M. G., Tanaka, M., et al. 2010, *MNRAS*, 403, 2063
- Frieman, J. A., Turner, M. S., & Huterer, D. 2008, *ARA&A*, 46, 385
- Frigo, M. & Johnson, S. G. 2005, *Proceedings of the IEEE*, 93, 216, special issue on “Program Generation, Optimization, and Platform Adaptation”
- Fruscione, A., McDowell, J. C., Allen, G. E., et al. 2006, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 6270, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*
- Georgakakis, A., Nandra, K., Laird, E. S., Aird, J., & Trichas, M. 2008, *MNRAS*, 388, 1205
- Giacconi, R. 2013, *Mem. Soc. Astron. Italiana*, 84, 472
- Giacconi, R., Gursky, H., Paolini, F. R., & Rossi, B. B. 1962, *Physical Review Letters*, 9, 439
- Gilli, R., Comastri, A., Brunetti, G., & Setti, G. 1999, *New A*, 4, 45

- Gilli, R., Comastri, A., & Hasinger, G. 2007, *A&A*, 463, 79
- Haardt, F., Maraschi, L., & Ghisellini, G. 1994, *ApJ*, 432, L95
- Hasinger, G., Cappelluti, N., Brunner, H., et al. 2007, *ApJS*, 172, 29
- Hasinger, G., Miyaji, T., & Schmidt, M. 2005, *A&A*, 441, 417
- Heckman, T. M. & Best, P. N. 2014, *ARA&A*, 52, 589
- Helgason, K., Cappelluti, N., Hasinger, G., Kashlinsky, A., & Ricotti, M. 2014, *ApJ*, 785, 38
- Henley, D. B. & Shelton, R. L. 2013, *ApJ*, 773, 92
- Hickox, R. C., Jones, C., Forman, W. R., et al. 2009, *ApJ*, 696, 891
- Hickox, R. C. & Markevitch, M. 2006, *ApJ*, 645, 95
- Hickox, R. C. & Markevitch, M. 2007, *ApJ*, 661, L117
- Hill, G. J., Gebhardt, K., Komatsu, E., & MacQueen, P. J. 2004, in *American Institute of Physics Conference Series*, Vol. 743, *The New Cosmology: Conference on Strings and Cosmology*, ed. R. E. Allen, D. V. Nanopoulos, & C. N. Pope, 224–233
- Hiroi, K., Ueda, Y., Akiyama, M., & Watson, M. G. 2012, *ApJ*, 758, 49
- Hirschmann, M., Somerville, R. S., Naab, T., & Burkert, A. 2012, *MNRAS*, 426, 237
- Hogg, D. W. 1999, *ArXiv Astrophysics e-prints*, 9905116
- Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, *ApJS*, 163, 1
- Hütsi, G. 2006, *A&A*, 449, 891
- Hütsi, G., Gilfanov, M., Kolodzig, A., & Sunyaev, R. 2014, *A&A*, 572, A28
- Hütsi, G., Gilfanov, M., & Sunyaev, R. 2012, *A&A*, 547, A21
- Jeong, D. & Komatsu, E. 2009, *ApJ*, 691, 569
- Kaiser, N. 1987, *MNRAS*, 227, 1
- Kalberla, P. M. W., Burton, W. B., Hartmann, D., et al. 2005, *A&A*, 440, 775

- Kashlinsky, A., Arendt, R. G., Mather, J., & Moseley, S. H. 2005, *Nature*, 438, 45
- Kembhavi, A. K. & Narlikar, J. V. 2000, *Physics Today*, 53, 60
- Kenter, A., Murray, S. S., Forman, W. R., et al. 2005, *ApJS*, 161, 9
- Kim, M., Wilkes, B. J., Kim, D.-W., et al. 2007, *ApJ*, 659, 29
- Kolodzig, A., Gilfanov, M., Hütsi, G., & Sunyaev, R. 2013a, *A&A*, 558, A90
- Kolodzig, A., Gilfanov, M., Sunyaev, R., Sazonov, S., & Brusa, M. 2013b, *A&A*, 558, A89
- Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, *ApJS*, 192, 18
- Koutoulidis, L., Plionis, M., Georgantopoulos, I., & Fanidakis, N. 2013, *MNRAS*, 428, 1382
- Krumpe, M., Lamer, G., Markowitz, A., & Corral, A. 2010a, *ApJ*, 725, 2444
- Krumpe, M., Miyaji, T., & Coil, A. L. 2010b, *ApJ*, 713, 558
- Krumpe, M., Miyaji, T., & Coil, A. L. 2014, in *Multifrequency Behaviour of High Energy Cosmic Sources*, 71–78
- Krumpe, M., Miyaji, T., Coil, A. L., & Aceves, H. 2012, *ApJ*, 746, 1
- La Franca, F., Fiore, F., Comastri, A., et al. 2005, *ApJ*, 635, 864
- Lehmer, B. D., Xue, Y. Q., Brandt, W. N., et al. 2012, *ApJ*, 752, 46
- Lewis, A. & Bridle, S. 2002, *Phys. Rev. D*, 66, 103511
- Lumb, D. H., Warwick, R. S., Page, M., & De Luca, A. 2002, *A&A*, 389, 93
- Mendez, A. J., Coil, A. L., Aird, J., et al. 2013, *ApJ*, 770, 40
- Merloni, A., Predehl, P., Becker, W., et al. 2012, *ArXiv e-prints*, 1209.3114
- Miyaji, T., Hasinger, G., Salvato, M., et al. 2015, *ArXiv e-prints*, 1503.00056
- Miyaji, T., Hasinger, G., & Schmidt, M. 2000, *A&A*, 353, 25
- Miyaji, T., Krumpe, M., Coil, A. L., & Aceves, H. 2011, *ApJ*, 726, 83

- Moretti, A., Campana, S., Lazzati, D., & Tagliaferri, G. 2003, *ApJ*, 588, 696
- Moretti, A., Vattakunnel, S., Tozzi, P., et al. 2012, *A&A*, 548, A87
- Mountrichas, G. & Georgakakis, A. 2012, *MNRAS*, 420, 514
- Mountrichas, G., Georgakakis, A., Finoguenov, A., et al. 2013, *MNRAS*, 430, 661
- Murray, S. S., Kenter, A., Forman, W. R., et al. 2005, *ApJS*, 161, 1
- Mushotzky, R. F., Done, C., & Pounds, K. A. 1993, *ARA&A*, 31, 717
- Nandra, K., O'Neill, P. M., George, I. M., & Reeves, J. N. 2007, *MNRAS*, 382, 194
- Padmanabhan, N., Xu, X., Eisenstein, D. J., et al. 2012, *MNRAS*, 427, 2132
- Peebles, P. J. E. 1980, *The large-scale structure of the universe* (Princeton University Press)
- Peebles, P. J. E., Daly, R. A., & Juszkievicz, R. 1989, *ApJ*, 347, 563
- Peebles, P. J. E. & Yu, J. T. 1970, *ApJ*, 162, 815
- Perinati, E., Tenzer, C., Santangelo, A., et al. 2012, *Experimental Astronomy*, 33, 39
- Pierre, M., Pacaud, F., Juin, J. B., et al. 2011, *MNRAS*, 414, 1732
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, *ArXiv e-prints*, 1502.01589
- Predehl, P., Andritschke, R., Böhringer, H., et al. 2010, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 7732, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
- Pringle, J. E. 1981, *ARA&A*, 19, 137
- Prokopenko, I. G. & Gilfanov, M. R. 2009, *Astronomy Letters*, 35, 294
- Reeves, J. N. & Turner, M. J. L. 2000, *MNRAS*, 316, 234
- Reynolds, M. T., Reis, R. C., Miller, J. M., Cackett, E. M., & Degenaar, N. 2014, *MNRAS*, 441, 3656
- Richards, G. T., Fan, X., Newberg, H. J., et al. 2002, *AJ*, 123, 2945

- Richardson, J., Chatterjee, S., Zheng, Z., Myers, A. D., & Hickox, R. 2013, *ApJ*, 774, 143
- Roncarelli, M., Cappelluti, N., Borgani, S., Branchini, E., & Moscardini, L. 2012, *MNRAS*, 424, 1012
- Roncarelli, M., Moscardini, L., Borgani, S., & Dolag, K. 2007, *MNRAS*, 378, 1259
- Rosati, P., Borgani, S., & Norman, C. 2002, *ARA&A*, 40, 539
- Saglia, R. P., Tonry, J. L., Bender, R., et al. 2012, *ApJ*, 746, 128
- Salvato, M., Ilbert, O., Hasinger, G., et al. 2011, *ApJ*, 742, 61
- Sawangwit, U., Shanks, T., Croom, S. M., et al. 2012, *MNRAS*, 420, 1916
- Schlegel, D., Abdalla, F., Abraham, T., et al. 2011, *ArXiv e-prints*, 1106.1706
- Shakura, N. I. & Sunyaev, R. A. 1973, *A&A*, 24, 337
- Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *MNRAS*, 323, 1
- Shu, X. W., Yaqoob, T., & Wang, J. X. 2010, *ApJS*, 187, 581
- Silverman, J. D., Green, P. J., Barkhouse, W. A., et al. 2008, *ApJ*, 679, 118
- Śliwa, W., Soltan, A. M., & Freyberg, M. J. 2001, *A&A*, 380, 397
- Slosar, A., Iršič, V., Kirkby, D., et al. 2013, *J. Cosmology Astropart. Phys.*, 4, 26
- Snowden, S. L., Egger, R., Freyberg, M. J., et al. 1997, *ApJ*, 485, 125
- Starikova, S., Cool, R., Eisenstein, D., et al. 2011, *ApJ*, 741, 15
- Sun, M. 2012, *New Journal of Physics*, 14, 045004
- Sun, M., Voit, G. M., Donahue, M., et al. 2009, *ApJ*, 693, 1142
- Sunyaev, R. A. & Zeldovich, Y. B. 1970, *Ap&SS*, 7, 3
- Tanaka, Y. 2013, *Mem. Soc. Astron. Italiana*, 84, 485
- Tegmark, M., Eisenstein, D. J., Strauss, M. A., et al. 2006, *Phys. Rev. D*, 74, 123507
- Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, *ApJ*, 480, 22

- Tenzer, C., Warth, G., Kendziorra, E., & Santangelo, A. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7742, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series
- Tozzi, P., Gilli, R., Mainieri, V., et al. 2006, *A&A*, 451, 457
- Truemper, J. 1993, *Science*, 260, 1769
- Ueda, Y., Akiyama, M., Hasinger, G., Miyaji, T., & Watson, M. G. 2014, *ApJ*, 786, 104
- Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, *ApJ*, 598, 886
- Urry, C. M. & Padovani, P. 1995, *PASP*, 107, 803
- Ursino, E., Branchini, E., Galeazzi, M., et al. 2011, *MNRAS*, 414, 2970
- Ursino, E., Galeazzi, M., & Haffenberger, K. 2014, *ApJ*, 789, 55
- Vajgel, B., Jones, C., Lopes, P. A. A., et al. 2014, *ArXiv e-prints*, 1409.6822
- Vikhlinin, A., Kravtsov, A., Forman, W., et al. 2006, *ApJ*, 640, 691
- Voges, W., Aschenbach, B., Boller, T., et al. 1999, *VizieR Online Data Catalog*, 9010, 0
- Wall, J. V. & Jenkins, C. R. 2012, *Practical Statistics for Astronomers*, 2nd edn. (Cambridge University Press)
- Watson, M. G., Schröder, A. C., Fyfe, D., et al. 2009, *A&A*, 493, 339
- Weinberg, D. H., Mortonson, M. J., Eisenstein, D. J., et al. 2013, *Phys. Rep.*, 530, 87
- Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, *ApJS*, 195, 10
- Yang, Q.-X., Xie, F.-G., Yuan, F., et al. 2015, *MNRAS*, 447, 1692
- Zwicky, F. 1933, *Helvetica Physica Acta*, 6, 110

Acknowledgment

Thanks, Danke, Gracias and Grazie to everyone who made this scientific adventure possible for me.

First of all I like to thank Marat Gilfanov for taken me in his supervision, for teaching me scientific working on the bleeding edge of science and for pushing me always to my limits. Many thanks also to Rashid Sunyaev, who he has an amazing ability to motivate and push me in order to reach my best performance, and for his constant support despite the odds. After almost one year into my PhD I obtained my third major collaborator with Gert Hütsi. If you would think of Marat and Rashid as my scientific father and grandfather, than Gert would probably be my big brother. Without him a big part of my PhD would have not been possible. Thank you all for sharing your scientific expertise with me and for the countless helpful discussions and advises. All of them are with no doubt one the smartest persons I have met and I am more than pleased to have worked with such excellent scientists.

Doing my PhD in Garching, which such a high density of astrophysicists, gave me the possibility to encounter many scientists with great knowledge in specific fields and I appreciated their open attitude to share their knowledge with me. First of all I need to thank Mirko Krumpe, who not just gave me great insides into large-scale structure studies with AGN but also into the scientific world in general. When it comes to AGN clustering I also received very useful consulting by Viola Allevato and Antonis Georgakakis, which I am thankful for. Since eROSITA is a major part of my PhD, I consider myself lucky to have received information out of first hand from the neighboring institute MPE despite the official limiting policies. Marcella Brusa became more by accident but due to her helpfulness my first source of information about eROSITA, which eventually led to a very fruitful collaboration for my first publication, and I am very happy about that. I am also very thankful to Andrea Merloni (the scientific PI of eROSITA), who became a great supporter of my scientific projects, which undoubtedly made my work in many ways much easier. And then there are of course several others, like Peter Pedrehl (the PI of eROSITA), Mara Salvato or Thomas Boller, who gave me valuable information about the eROSITA

mission and beyond.

I encountered many ambitious PhD students and post-docs during my PhD, which were always very helpful and supportive, and this spirit is also one of the reasons why I like such much working in science. Robert Andrassy was my longest office mate and a grateful receiver for all the little and stupid questions I had in my mind. Thank you for taken your time and sharing your scientific wisdom with me. I think one day he will become a very kind professor. Starting with analyzing real X-ray data, I was lucky to have Mike Anderson around. His expertise were very helpful and I also very much enjoyed the scientific discussion with him in general, since he is a pretty smart guy with a great career ahead of him. Coding was my major task in the PhD and with the help and advises of the coding genius Philipp Edelmann it become a lot easier. I think that he should get paid a higher salary for his great and voluntary IT-support. When I entered the large and chaotic world of science at MPA, Lindy became quickly my anchor point. The transition into this world was nevertheless though, but her company made it a lot more smoother for me.

The path to my PhD was paved by tree important characters. It all started with my astrophysical lectures by Lutz Wisotzki back at my home university. He strongly encouraged me to pursue my passion in astrophysics. My first real encounter with scientific data was made with the help of Maria Diaz-Trigo. The traineeship with her was not just a great experience in general but also a major boost of my scientific career. I am also happy that Maria continued supporting me and my career after my traineeship. Last but not least, i was very lucky to meet Axel Schwope, who helped me to improve and broaden my scientific skills. This accumulated in me doing my master thesis (Diplomarbeit) with him on a very interesting topic and I very much enjoyed his supervision. I owe all three of them a great deal of gratitude.

Besides the scientific support I had also plenty of support by friends, colleagues and acquaintances. If you are actually made it this far and reading this now, than your are most likely one of them! Thank you for making my life so much more enjoyable.

In the end I have a very special thanks to my parents, which gave me the opportunity of pursuing my dreams with their full support and love. *Vielen Dank Mutti und Vati für eure Unterstützung und Liebe. Ohne euch hätte ich es sicher nicht so weit gebracht und hätte nie mein Potential soweit ausreizen können. Ihr braucht nicht stolz auf mich sein, sondern ihr könnt stolz auf euch sein!*

